A Delta Smelt Individual-Based Life Cycle Model in the R Statistical Environment

William Smith
July 15, 2022

Executive summary

The delta smelt individual-based life cycle model in R (IBMR) simulates four processes of the San Francisco Estuary's (SFE) delta smelt population: reproduction, movement, growth, and mortality. The model provides a mechanistic description of the life cycle, synthesizing the most recent life history information and environmental data measured at spatiotemporal scales relevant to delta smelt ecology and management. Environmental conditions are represented by temperature, turbidity, Old and Middle River flow, and prey density.

The IBMR is a modification of the IBM in Fortran (DSIBM), presented by Rose et al. (2013a and 2013b). By design, many features of IBMR are identical to Rose's original DSIBM, but IBMR was developed to be more accessible by using a more common statistical program and providing a set of open-source code that can be accessed and run by any R-user.

IBMR was calibrated to abundances and growth rates estimated for the wild delta smelt population. The model is suitable for assessing the population growth potential of delta smelt given changes to abiotic and biotic conditions in the SFE.

Contents Page

Modifications in the current IBMR v3

Several variations of a delta smelt individual-based life cycle model have been developed since Rose et al. (2013a) published their original model (DSIBM). Since 2018, at least three independent groups have modified or updated the DSIBM. A summary of the version history of the model is included in Table 1. In preparation for the CSAMP Delta Smelt Structured Decision Making Process in 2021, the IBMR was further modified (IBMR v3). All data were updated for years 1995 through at least 2014; the bioenergetics model was updated to include the 12 prey types developed by Kimmerer for DSIBM v2 (Appendix A), and reproductive parameters were updated based on the most recent published information (Damon et al. 2016). IBMR v3 retained the monthly time step of IBMR v1 and v2. The spawning model was updated to include a third temperature-dependent spawning event in April, allowing up to 3 batches of eggs per spawner, conditional on temperature and survival (LaCava et al. 2015; Kurobe et al. 2016). Finally, the consumption component of the bioenergetics model was updated to include the turbidity effect on foraging rates documented by Hasenbein et al. (2016) and for other fish species occupying a similar ecological niche to delta smelt (Pangle et al. 2012).

Data

Five physical and biological variables, representing observed Delta conditions during 1995-2014, drive simulated population dynamics: prey density, Old and Middle River flow, delta smelt distribution, water temperature, and Secchi depth. All data used to summarize IBMR variables may be found online (https://github.com/CSAMP). All variables except Old and Middle River flow had dimensions year *y*, month *m*, and spatial strata *s* (Table 2). Old and Middle River flow was a *y* x *m* matrix, and prey densities included a fourth dimension *p* indexing prey type.

Prey density. Prey density estimates PD_{ymsp} for 12 zooplankton prey types *p* and years 1989-2015 were developed by Wim Kimmerer (Appendix A). Daily log prey carbon densities (mgC m^{-3}), estimated from positive catches, and daily proportion zero catch were summarized from the Interagency Ecological Program's zooplankton survey and the CA Department of Fish and Wildlife (CDFW) 20-mm survey. Monthly means of the daily values were used to simulate prey densities.

Old and Middle River flow. OMR_{ym} was the monthly average of the daily sum of tidally filtered flows (ft^3 s^{-1}) from two adjacent channels of the San Joaquin River basin, Old and Middle Rivers (Fig. 2). *OMR* data were available from US Geological Survey (USGS) streamflow databases (https://waterdata.usgs.gov). Much of the daily streamflow gauge data was missing. If data for one river was missing, daily *OMR* was predicted from a linear model of the flow in the remaining river, and if data for both rivers was missing, the model proposed by Andrews et al. (2016) was used to estimate missing *OMR* from San Joaquin River flows and exports from South Delta water

diversions. Daily San Joaquin River flow and export data were available from the Dayflow database (https://data.cnra.ca.gov/dataset/dayflow).

Fish distribution. Observed delta smelt distributions DS_{yms} (proportion in each stratum) were developed from 20-mm, Midwater Trawl, and Spring Kodiak Surveys. Townet Survey data were not used, because tow volume data were unavailable before 2003, use of Townet data would only leverage one additional month of distribution information (August), and in most year-months with a concurrent 20-mm Survey sample, delta smelt were observed in fewer spatial strata in the Townet Survey (in 7 of 13 years). Fish survey data were made available online by CDFW (ftp://ftp.dfg.ca.gov/). Monthly observed catch densities (catch/volume sampled) were assigned to the 12 spatial strata, and month-strata mean observed densities were expanded to population abundance by multiplying by strata volumes. Strata volumes were estimated by Derek Hilts (USFWS; Bay-Delta Office) using DSM2. Estimated population abundances in each spatial stratum were then converted to proportions of the total abundance, which were interpreted as observed occupancy probabilities.

Surveys were not completed in all year-months, and in some months, no fish were observed (mostly February, March, and August). Distributions from the prior month were assumed when observations were missing. The fish distribution data was characterized by an increasing level of zero-inflation, resulting in zero observed occupancy of many spatial strata towards the end of the time series. 2014 was selected as the terminal cohort of observed spatial distributions (last observed in April 2015) in order to mitigate this issue.

Delta smelt densities in the South Delta spatial stratum were particularly difficult to characterize, with many instances in which entrained delta smelt were observed at the pumping facilities, but not in the South Delta fish surveys. The Delta Smelt Life Cycle Model (Smith et al. 2021) included estimates of the fraction of the population that entered the South Delta and were eliminated from the population (entrained). These estimates of proportional entrainment loss were developed by integrating multiple data sources, including salvage or observations of delta smelt at the pumping facilities. Life Cycle Model estimates of proportional entrainment were interpreted as the fraction of the population in the South Delta, and adopted as IBMR estimates of the South Delta distribution of delta smelt.

An alternative treatment of the South Delta distributions used the fish surveys to estimate proportion in the South Delta, like other strata; however, delta smelt densities in the South Delta stratum were depleted by entrainment, resulting in a negative bias and underestimation of the proportion of the population in this region of the Delta. Further, on many occasions, no fish were observed in the South Delta surveys, but one or more fish were salvaged, indicating that South Delta survey data were more zero-inflated than salvage data.

The conceptual model for the entrainment process is that under certain conditions, a fraction of the population is induced to occupy a portion of the Lower San Joaquin River, where they become

vulnerable to advective flows (*OMR*) into the South Delta. Over a period of 1-2 weeks, the group of fish that moved into the South Delta becomes depleted by water exports until no or few fish remain. The model of delta smelt distributions did not account for the conditions leading fish to occupy the South Delta; it only accounted for the fraction of fish that were originally observed there.

Temperature. Water temperature data $Temp_{\text{yms}}$ (Celsius) for years 1990-2010 were summarized from DSM2 hydrodynamic simulations by Derek Hilts (Fig. 2). The terminal year of the DSM2 *Temp* dataset, 2010, limited the number of years available for the IBMR by at least 4 years. A second set of water temperature data was therefore summarized from all available online data collected by the CDFW and FWS during Delta fish monitoring programs (ftp://ftp.dfg.ca.gov/; https://www.fws.gov/lodi/). The second water temperature dataset spanned years 1959-2020, but prior to 2011, data for some year-month-strata combinations were missing or sparse, with only a single sample.

The water temperatures that would have been simulated by DSM2 for missing years, 2011-2014, were predicted using a general linear model of DSM2 monthly means $Temp_{\text{yms}}$ as a function of season, spatial strata, and monthly mean temperatures measured by fish monitoring programs $\widehat{Temp}_{\text{yms}}$. The best model of $Temp_{\text{yms}}$ was identified using backwards selection, starting with a full model, having an effect for each stratum and season, and eliminating non-significant spatial effects (acceptance level < 0.05), one coefficient at a time, before eliminating non-significant seasonal effects. The model was fit to data from years 1990-2010, and used to predict $Temp_{\text{yms}}$ from measured $\widehat{Temp}_{\text{yms}}$ for years 2011-2014. See Appendix B for details of the model to predict $Temp_{\text{yms}}$.

Secchi depth. Secchi depth data $Secchi_{\text{yms}}$ (cm) were used as an index of turbidity. Like $\widehat{Temp}_{\text{yms}}$, Secchi depths for years 1959-2020 were summarized from all CDFW and FWS databases available online (Fig. 2). Means for each year-month-stratum combination were summarized, but data were not available for all strata in all year-months. Missing data were estimated from general linear models of the remaining Secchi data in other spatial strata. The best general linear model for each stratum was selected using backwards selection, beginning with the full model, having a separate coefficient for each spatial stratum, and eliminating non-significant coefficients, one at a time. See Appendix B for details of the model to predict missing $Secchi_{\text{yms}}$.

Model

The IBMR v3 (Fig. 3) includes cohorts spawned from 1995 to 2014. A cohort year begins at the beginning of simulated spawning in February and ends the following January. Reproduction, movement among 12 spatial strata, growth, and mortality of a closed population are modeled (Fig.

1). The simulated population in one year depends on the characteristics of the simulated population in the previous year; therefore, it is important to establish stable length and spatial distributions before simulating the 1995-2014 time series. IBMR establishes stable length and spatial distributions by looping through a representative sequence of years before simulating the 1995-2014 time series. The first (four) loops are discarded as a burn-in. This results in median fork lengths of 74 mm in the initial month of the simulation, February 1995.

Starting conditions. After the super-individual approach of the DSIBM, a super-adult approach was developed for IBMR. At the end of each simulated year (end of January), the adult population was reset to 200 super-adults. Population growth each was then then number of adults after 12 months, the following January, divided by 200. The super-individual approach was developed to speed up computation time by avoiding very large numbers of simulated individuals while also avoiding 'crashes' when the simulated population declines to zero.

The model is initialized with adult spawners in February and assumed weights of adult delta smelt, generated from a lognormal distribution. Weights are converted to lengths using the length-weight equation derived by Kimmerer et al. (2005),

$$(1) \qquad L_{\mathrm{iym}} = \sqrt[3.82]{\frac{W_{\mathrm{iym}}}{1.83e^{-6}}},$$

where W_{iym} is weight in grams and L_{iym} is fork length in millimeters of individual *i* in year *y* and month *m* of the simulation.

Reproduction. Four reproductive sub-processes are modeled: maturation, sex ratio, fecundity, and egg to larval survival. Simulated delta smelt reproduction occurs during the months of February, March, and April. During spawning months, maturity is stochastically assigned using a Bernoulli distribution and a length-based, logistic regression model of maturation probability

$$(2) \qquad mature_{\mathrm{iym}} \sim \mathrm{Bernoulli}\left(1/\left(1 + e^{-0.1*(L_{\mathrm{iym}}-60)}\right)\right),$$

where 60 mm represented an assumed length at 50% maturity and 0.1 represents an assumed slope coefficient. With this maturation model, the probability of maturity increases from approximately 0.05 at 48 mm to 0.95 at 72 mm (Fig. 4). This is in contrast to the DSIBM, which used a knife-edged maturation model, in which individuals less than 60mm FL had no probability of maturation and spawning.

Sex is also stochastically assigned to simulated fish using a Bernoulli distribution, with a 0.48 probability of female assignment, based on Spring Kodiak Trawl samples of adult delta smelt during the spawning season

$$(3) \qquad female_{\mathrm{i}} \sim \mathrm{Bernoulli}(0.48).$$

Water temperature is assumed to affect several components of the reproductive and early life history of delta smelt, including the number of potential batch spawns, egg to larval survival, and length at first feeding (Rose et al. 2013a). Mature females spawn up to three batches of eggs, with one potential batch in both February and March and an April batch if an individual occupies a stratum with *Temp* < 17.9°C, based on the maximum observed spawning temperatures noted by Damon et al. (2016). Population egg production is based on the distribution of *L*, the length-based fecundity model estimated by Damon et al. (2016) (Fig. 4), and the simulated abundance of mature females (adults; AD)in each year-month-stratum $n_{\mathrm{AD_{yms}}}|female, mature$. Total egg production in each year-month-stratum $n_{\mathrm{Egg_{yms}}}$ of the spawning season is the sum of fecundity of mature females in that stratum

(4) $$n_{\mathrm{Egg_{yms}}} = \sum_{j=1}^{n_{\mathrm{AD_{yms}}}|female,mature} 0.0183 * L_{\mathrm{jym}}^{2.7123}.$$

Simulated eggs transition to feeding larvae at the end of the month of spawning. Survival of eggs to the larval life stage $S_{\mathrm{LRV_{ym}}}$ (Fig. 5) is the product of proportion hatching $P_{\mathrm{Hatch_s}}$ and egg to larval survival $S_{\mathrm{YS_{yms}}}$,

(5) $$S_{\mathrm{LRV_{yms}}} = P_{\mathrm{Hatch_s}} * S_{\mathrm{YS_{yms}}}, \text{ where}$$

(6) $$P_{\mathrm{Hatch_s}} = -2.35 + 0.45 * Temp_{\mathrm{yms}} + 0.016 * Temp_{\mathrm{yms}}^2 \text{ and}$$

(7) $$S_{\mathrm{YS_{yms}}} = e^{-0.035*(-10.1-1.5*Temp_{\mathrm{yms}})}.$$

Parameters defining the $P_{\mathrm{Hatch_s}}$ and $S_{\mathrm{YS_{yms}}}$ models (Eq. 6-7) were given by Bennett (2005) (Bennett's Fig. 10) and included temperature effects. Number of days as a yolk-sac larvae $(-10.1 - 1.5 * Temp_{\mathrm{yms}})$ is multiplied by daily yolk-sac larval mortality of 0.035 (Eq. 7), estimated by Rose et al. (2013a), to predict $S_{\mathrm{YS_{yms}}}$.

Initial simulated lengths L_{LRV} for larvae at the end of their first month, before entering the feeding population in the bioenergetics model, are calculated

(8) $$L_{\mathrm{LRV_i}} = 5.92 - 0.05 * Temp_{\mathrm{yms}},$$

and $L_{\mathrm{LRV_i}}$ are converted to weights using the larval length-weight equation derived by Kimmerer et al. (2005),

(9) $$W_{\mathrm{iym}} = 5e^{-6} * L_{\mathrm{iym}}^3.$$

Movement. A Eulerian random walk model was developed to simulate delta smelt movement among spatial strata that approximates the observed spatial distributions of the population. Fish are randomly assigned to one of n_{STR} spatial strata based on observed spatial distributions DS_{yms}

and a set of rules for movement. Starting strata for each individual *i* at the end of January 1995 are categorically distributed, with probabilities equal to DS_{yms}

(10) $strata_{\mathrm{iym}} \sim \mathrm{Categorical}\left(DS_{\mathrm{yms}}\right).$

Starting strata for newly spawned fish are assigned to the parental spatial strata. In each subsequent month, categorical distribution probabilities to randomly assign strata are the product of observed spatial distributions and a set of rules for month-to-month movement that allow residence in a stratum, movement between adjacent strata, or movement between strata separated by only a single stratum. The movement rules prevent simulated movement to more distant strata (e.g., month-to-month movement between Yolo and Suisun Bay strata is not modeled). Rules for movement between stratum s and s′ are represented by an n_{STR} x n_{STR} matrix $R_{\mathrm{ss}\prime}$ of 1s and 0s, where 1s permit simulated movement between a pair of spatial strata and 0s prevent simulated movement

(11) $strata_{\mathrm{iym}} \sim \mathrm{Categorical}\left(DS_{\mathrm{yms}} * R_{\mathrm{s}(1:n_{\mathrm{STR}})}\right).$

Growth

The bioenergetics growth model described by Rose et al. (2013) was a system of equations estimating daily delta smelt growth in body mass as a function of rates of consumption C_{iym}, metabolism R_{iym}, egestion F_{iym}, excretion U_{iym}, activity SDA_{iym} and spawning losses Sp_{iym} in for individual *i* in year *y* and month *m* (Eqs. 12-24). In this application, daily growth increments were scaled to monthly increments by multiplying by the number of days in each month. A set of bioenergetics model coefficients, specific to each life-stage *l* to model each rate were listed in Rose et al. (2013a) (Fig. 7); in the notation below, these fixed coefficients are underlined to distinguish them from dynamic quantities that may vary by time period.

(12) $W_{\mathrm{iy(m+1)}} = W_{\mathrm{iym}} * \left(1 + n.days_m * \frac{ep_{\mathrm{iym}}}{4814} * \left(C_{\mathrm{iym}} - R_{\mathrm{iym}} - F_{\mathrm{iym}} - U_{\mathrm{iym}} - SDA_{\mathrm{iym}}\right)\right) - Sp_{\mathrm{iym}}$, where

(13) $R_{\mathrm{iym}} = \underline{ar_{\mathrm{l}}} * W_{\mathrm{iym}}^{\underline{br_{\mathrm{l}}}} * e^{\underline{RQ_{\mathrm{l}}} * Temp_{\mathrm{yms}}}$,

(14) $F_{\mathrm{iym}} = \underline{Fa_{\mathrm{l}}} * C_{\mathrm{iym}}$,

(15) $U_{\mathrm{iym}} = \underline{Ua_{\mathrm{l}}} * \left(C_{\mathrm{iym}} - F_{\mathrm{iym}}\right)$,

(16) $SDA_{\mathrm{iym}} = \underline{Sd_{\mathrm{l}}} * \left(C_{\mathrm{iym}} - F_{\mathrm{iym}}\right)$, and

(17) $Sp_{\mathrm{iym}} = 0.15 * W_{\mathrm{iym}}$

The conversion of prey to delta smelt biomass was expected to be less efficient for *Limnoithona* prey because of its lower energy density ed_{p}. The lower ed_{p} of *Limnoithona* was accounted by

adjusting the efficiency at which simulated consumption was converted to delta smelt weight, represented by the ratio $ep_{iym}/4814$ (Eq. 12). ep_{iym} was the energy density of prey consumed, reduced by the fraction of consumed energy corresponding to *Limnoithona* (Eq. 18 and 19), and 4,814 J/g was the energy density of delta smelt. The energy density of *Limnoithona* (1,813 J/g) was assumed to be 30% less than that of all other prey items (2,590 J/g).

(18) $$ep_{\mathrm{iym}} = 1813 * Limno_{\mathrm{iym}} + 2590 * \left(1 - Limno_{\mathrm{iym}}\right), \text{ where}$$

(19) $$Limno_{\mathrm{iym}} = \frac{1813 * Cmax_{\mathrm{iym}} * \left(\frac{\frac{PD_{\mathrm{ym(LImno)}} * \underline{V_{\mathrm{(Limno)l}}}}{\underline{K_{\mathrm{(Limno)l}}}}}{\sum_{r=1}^{12} \frac{PD_{\mathrm{ymr}} * \underline{V_{\mathrm{rl}}}}{\underline{K_{\mathrm{rl}}}}} \right)}{\sum_{q=1}^{12} \underline{ed_{\mathrm{q}}} * Cmax_{\mathrm{iym}} * \left(\frac{\frac{PD_{\mathrm{ymq}} * \underline{V_{\mathrm{ql}}}}{\underline{K_{\mathrm{ql}}}}}{\sum_{r=1}^{12} \frac{PD_{\mathrm{ymr}} * \underline{V_{\mathrm{rl}}}}{\underline{K_{\mathrm{rl}}}}} \right)}$$

where PD_{ymp} was the prey density of prey type p.

The maximum possible consumption rate $C_{\mathrm{max_{iym}}}$ was a measure of potential foraging rate, expressed as a proportion of body weight per day (Eqs. 20 and 21). Foraging arena theory suggests that fish reduce their time spent foraging to mitigate perceived risk of mortality, at the expense of forgone foraging and growth. Two environmental constraints on delta smelt foraging were considered: temperature $Temp_{\mathrm{ym}}$ effects ($KA_{\mathrm{iym}} * KB_{\mathrm{iym}}$) and turbidity $Turb_{\mathrm{iym}}$ effects (KT_{iym}). Relationships between *Temp*, *C*, and *R* are shown in Fig. 6.

(20) $$C_{\mathrm{iym}} = C_{\mathrm{max_{iym}}} * \sum_{q=1}^{12} \left(\frac{\frac{PD_{\mathrm{ymp}} * \underline{V_{\mathrm{pl}}}}{\underline{K_{\mathrm{ql}}}}}{\sum_{r=1}^{12} \frac{PD_{\mathrm{ymp}} * \underline{V_{\mathrm{pl}}}}{\underline{K_{\mathrm{rl}}}}} \right), \text{ where}$$

(21) $$C_{\mathrm{max_{iym}}} = \underline{ac_{\mathrm{l}}} * W_{\mathrm{iym}}^{\underline{bc_{\mathrm{l}}}} * KA_{\mathrm{yms}} * KB_{\mathrm{yms}} * KT_{\mathrm{yms}}$$

Rose et al. (2013) assumed a *Temp*-C_{max} model for delta smelt (KA_{yms} and KB_{yms}; Eq. 12 and13) that reduced foraging time as water temperatures increased above 23°C (Fig. 6).

(22) $$KA_{\mathrm{yms}} = \frac{\underline{CK1_{\mathrm{l}}} * e^{\frac{1}{\underline{T0_{\mathrm{l}}} - \underline{CQ_{\mathrm{l}}}} * ln\left(\frac{0.98 * \left(1 - \underline{CK1_{\mathrm{l}}}\right)}{0.02 * \underline{CK1_{\mathrm{l}}}}\right) * \left(Temp_{\mathrm{yms}} - \underline{CQ_{\mathrm{l}}}\right)}}{1 + \underline{CK1_{\mathrm{l}}} * \left(e^{\frac{1}{\underline{T0_{\mathrm{l}}} - \underline{CQ_{\mathrm{l}}}} * ln\left(\frac{0.98 * \left(1 - \underline{CK1_{\mathrm{l}}}\right)}{0.02 * \underline{CK1_{\mathrm{l}}}}\right) * \left(Temp_{\mathrm{yms}} - \underline{CQ_{\mathrm{l}}}\right)} - 1 \right)}$$

$$(23)\quad KB_{\mathrm{yms}} = \frac{\underline{CK4_{\mathrm{l}}} * e^{\frac{1}{\underline{TL_{\mathrm{l}}} - \underline{TM_{\mathrm{l}}}} * ln\left(\frac{0.98*\left(1-\underline{CK4_{\mathrm{l}}}\right)}{0.02*\underline{CK4_{\mathrm{l}}}}\right) * \left(\underline{TL_{\mathrm{l}}} - Temp_{\mathrm{yms}}\right)}}{1+\underline{CK4_{\mathrm{l}}} * \left(e^{\frac{1}{\underline{TL_{\mathrm{l}}} - \underline{TM_{\mathrm{l}}}} * ln\left(\frac{0.98*\left(1-\underline{CK4_{\mathrm{l}}}\right)}{0.02*\underline{CK4_{\mathrm{l}}}}\right) * \left(\underline{TL_{\mathrm{l}}} - Temp_{\mathrm{yms}}\right)} \right) - 1)}$$

Forage fish, such as delta smelt, typically show a decrease in foraging rates as turbidity declines and the perceived risk of being detected by a predator increases (Pangle et al. 2012). The risk of predation and changes in delta smelt behavior in clear water were documented by Ferrari et al. (2014), though rates of predation may have been biased high because smelt could not effectively evade predators in laboratory conditions. The relationship between delta smelt foraging rate and turbidity reported by Hasenbein et al. (2016) was approximated using a simple logistic model (Fig. 2), that increased from the lowest turbidities evaluated (5 NTU; 84 cm Secchi depth) to the turbidities associated with maximum foraging rate (25-80 NTU; 28-13 cm Secchi depth). Since turbidities greater than 80 NTU were rarely observed during the time period explored, foraging limitation at high turbidity was not modeled, i.e., using a dome-shaped double-logistic model. As turbidity declined, the effect of turbidity (KT_{yms}; Eq. 14) was assumed to reach some asymptotic minimum α_{FL}.

$$(24)\quad KT_{\mathrm{yms}} = \alpha_{FL} + (1 - \alpha_{FL})/\left(1 + e^{0.1*\left(Turb_{\mathrm{yms}} - 56.2\right)}\right)$$

Mortality. The survival probability of simulated juvenile to adult delta smelt $S_{\mathrm{JA_{iym}}}$ is the function of two competing sources of mortality, entrainment mortality *F* and natural mortality *M* (Fig. 8)

$$(25)\quad S_{\mathrm{JA_{iym}}} = e^{-\left(F_{\mathrm{iym}} + M_{\mathrm{iym}}\right)}.$$

F represents the effect of being drawn into the South Delta through the action of pumping, where fish may be subject to direct export from the system, predation in the South Delta, or otherwise isolation from the remainder of the population. *M* represents all other sources of mortality. Natural mortality in IBMR predominantly represents predation, as the effects of food and starvation are accounted with bioenergetics.

The survival state of each individual z_{iym} (1 or 0) is randomly assigned to each individual based on draws from a Bernoulli distribution with probability S_{iym}

$$(26)\quad z_{\mathrm{iym}} \sim \mathrm{Bernoulli}\left(S_{\mathrm{JA_{iym}}}\right).$$

Simulated fish may also die due to starvation, if their weight declines more than 15% during a one-month period.

F is applied to fish located in the South Delta stratum during December–June. A 'ramp' model of *F* as a function of *OMR* is used to calculate the level of *F* to apply (Fig. 8). If *OMR* is less than or

equal to the minimum threshold (-5,000 ft^3/s for larger fish and -3,500 ft^3/s for smaller fish), all fish occupying the South Delta experience entrainment mortality; at the simulated population level, the entrainment loss equals that estimated by the US Fish and Wildlife Service's Life Cycle Model (Smith et al. 2021). *F* declines linearly to zero from the LCM-based level, as *OMR* increases to 0 ft^3/s. Fish located outside of the South Delta are assigned $F = 0$; all fish are assigned $F = 0$ during July–November.

Rose et al. (2013b) defined a length-based *M* model $M_{\mathrm{iym}} = -0.034 + 0.165 * L_{\mathrm{iym}}^{-0.322}$. The same length-based model of *M* is assumed in IBMR v3, but it is scaled in proportion to the *Secchi*-$C_{\max}$ relationship (KT_{iym}; Eq. 24). It is assumed that the length-based model represented a maximum potential rate of *M*, and that delta smelt experience lower predation risk as turbidity increases. As turbidity increases and predation risk declines, delta smelt respond by increasing foraging rate (the *Secchi*-$C_{\max}$ relationship). In other words, this approach assumes that delta smelt limit foraging in proportion to the predation risk they perceive as turbidity varies. Although the experiments of Ferrari et al. (2014) demonstrated a mechanism relating lower turbidity to higher predation of delta smelt, these results may have been biased by the confined, artificial conditions of the experiment. Wild delta smelt are expected to experience less predation mortality compared to what Ferrari et al. (2014) documented in a lab setting. Under poor foraging conditions, represented by low turbidity, foraging is minimized ($KT_{\mathrm{iym}} = \alpha_{FL}$) and natural mortality is maximized at the rate predicted by the length-based *M* model. Under ideal conditions of higher turbidity, foraging is maximized ($KS_{\mathrm{iym}} = 1$) and natural mortality is minimized at the α fraction of delta smelt natural mortality predicted by the length-based *M* model.

$$M_{\mathrm{iym}} = \left(-0.034 + 0.165 * L_{\mathrm{iym}}^{-0.322}\right) * \left(1 - \alpha_{FL} + KT_{\mathrm{iym}}\right). \quad (27)$$

Population growth. Annual growth of the simulated delta smelt population is defined by spawner abundance (AB), or the number of fish in the simulated population just prior to the beginning of the spawning season in February. Annual population growth $\lambda_{\mathrm{AB_y}}$ is calculated

$$\lambda_{\mathrm{AB_y}} = \frac{AB_y}{AB_{y-1}}. \quad (28)$$

Model calibration and diagnostics. M and the maximum turbidity penalty on foraging α were calibrated so that the simulated delta smelt population reproduced certain qualities of the wild delta smelt population. *M* required calibration to generate the long-term patterns in delta smelt abundance observed across fish surveys. $\lambda_{\mathrm{AB_y}}$ were the product of reproductive, survival, and individual growth rates. IBMR survival rates (i.e., *M*) were calibrated to approximate the population growth rates observed in the wild population (Polansky et al. 2019). The geometric mean of observed population growth rates, calculated from 2000-2014 June abundance estimates was 0.8. To calibrate IBMR to observed abundances, M_{iym} were iteratively multiplied by a constant until the median geometric mean λ_{AB} among IBMR-simulated June abundance time series

was approximately 0.8. Note that June abundances were used in this calibration, in order to leverage one of the better observed delta smelt abundance datasets using the 20mm Survey.

The growth of simulated delta smelt depended on biotic and abiotic constraints on foraging. Biotic constraints were described by prey type and availability and the multispecies functional response (Eq. 20). Abiotic constraints were described by temperature and turbidity effects on $C_{\max}$ (Eq. 22-24), which are only partially informed by experimental results; thus, temperature and turbidity effects on $C_{\max}$ represent a critical uncertainty in IBMR dynamics. In order to assess whether IBMR reproduced the expected growth trajectory for delta smelt, simulated mean length at age was compared to two reference points: a von Bertalanffy growth model fit to lengths and ages observed in the wild population (Smith 2018) and mean lengths of delta smelt observed in February fish surveys. The maximum effect of low turbidity on consumption rates α was calibrated to observed February lengths by iteratively changing α to minimize the residual (observed lengths – IBMR-predicted lengths) sum of squares for years 2002-2014, when the initiation of Spring Kodiak Trawl sampling reduced the length-selectivity bias of the observed mean February lengths. Prior to 2002, mean observed lengths may have been biased low by the selectivity of the Midwater Trawl gear, which appears to decline for larger delta smelt (Mitchell et al. 2019).

After Rose et al. (2013b), a sensitivity analysis was performed to evaluate which data types IBMR dynamics were most sensitive to. Data that drove the simulated dynamics of the delta smelt population were Old and Middle River flow, water temperature, Secchi depth, prey density, and observed delta smelt spatial distributions. First, the years with the highest and lowest simulated January population growth rates were identified. Data in these years formed two alternate sets, representing conditions leading to high and low population growth. All years 1995-2014 of the IBMR simulation were run, holding each type of model input constant, one at a time, at the high or low sets, while allowing other data to vary at the observed values. IBMR sensitivity was indicated by the relative change in the geometric mean of $\lambda_{\mathrm{AB_y}}$. Using the model with all data varying at the observed 1995-2014 values as a baseline, greater distance of mean λ_{AB} from the baseline, when holding a particular data type or combination of data types constant indicated sensitivity in simulated IBMR dynamics.

A second set of sensitivity analyses were performed to explore model assumptions about fish distributions in the South Delta and the assumption that natural mortality varies as a function of Secchi depth. In alternate IBMR configurations, fish distributions in the South Delta were estimated from survey data, as for other spatial strata, and the Secchi depth mortality relationship (Eq. 26) was eliminated.

Results

A total of 300 1995-2014 delta smelt population time series were simulated in the IBMR, requiring approximately 3 hours of run time. Greater efficiency could be achieved using parallel processing.

Simulated proportional entrainment mortality increased at older life stages, as abundance declined (Fig. 9). All fish occupying the South Delta stratum were assumed to die of entrainment mortality when *OMR* was -5,000 cfs or more negative (Fig. 8), and no fish occupying the South Delta were assumed to die of entrainment mortality when *OMR* was greater than 0 cfs.

During calibration, the IBMR natural mortality model (Eq. 26) was scaled by a factor of 1.262, generating a median geometric mean λ_{AB}, among all simulated June abundance time series of 0.983 (Fig. 10). Though most observed abundances fell within the 95% interval of simulated abundances, modest lack of fit to the Life Cycle Model estimates (Smith et al. 2021) June abundance (recruitment) estimates occurred in two time periods. IBMR predicted somewhat higher recruitment than the Life Cycle Model during 2000-2005, and IBMR predicted lower recruitment than the Life Cycle Model during 2010-2011. Fit to other Life Cycle Model estimates of abundances was better than fit to observed June abundances.

Simulated lengths after one year of growth appeared to reach the lengths predicted by the von Bertalanffy growth model (Fig. 11). Although IBMR growth rates in the first four months were lower than predicted by the von Bertalanffy growth model, the von Bertalanffy predicted growth rates may be positively biased for younger fishes (Haddon 2001). A maximum effect of low turbidity on consumption rates $\alpha = 0.63$ maximized the fit to lengths measured in February (Fig. 12).

As indicated by the greatest change in mean λ_{AB} from the baseline, IBMR was most sensitive to the variation in prey density (*PD*) data and to combinations of *PD* and *Secchi*, fish distribution (*DS*), and *OMR* (Fig. 13). Mean λ_{AB} of the base model was 0.955, and maximum $\lambda_{\mathrm{AB}} = 2.15$ and minimum $\lambda_{\mathrm{AB}} = 0.397$ were simulated in years 1998 and 2004, respectively. When individual data types were fixed at the 1998 or 2004 values, while allowing all other data to vary at the observed 1995-2014 values, the highest mean λ_{AB} was simulated using 1998 levels of *PD* (mean $\lambda_{\mathrm{AB}} = 1.86$), and the lowest mean λ_{AB} was simulated using 2004 levels of *PD* (mean $\lambda_{\mathrm{AB}} = 0.511$). When combinations of two data types were set to the 1998 or 2004 values, the highest mean λ_{AB} was simulated using 1998 levels of *PD* and *DS* (mean $\lambda_{\mathrm{AB}} = 2.81$) and using 1998 levels of *PD* and *OMR* (mean $\lambda_{\mathrm{AB}} = 2.093$), and the lowest mean λ_{AB} was simulated using 2004 levels of *PD* and *DS* (mean $\lambda_{\mathrm{AB}} = 0.289$).

Alternate IBMR configurations, with either an alternate method to estimate South Delta *DS* or no turbidity relationship with natural mortality (Eq. 26) demonstrated similar model sensitivity to *PD*; however, models with South Delta *DS* estimated from fish surveys were less sensitive to *OMR*,

and models without a turbidity relationship with natural mortality were less sensitive to *Secchi* (Fig. C1-C2).

Recommended applications

The IBMR is a mechanistic representation of several critical processes of the delta smelt population, but it relies on major assumptions that determine exactly how some mechanisms operate. At the coarsest scale, the model may be used to characterize long-term annual population growth rates, and at a finer spatiotemporal scale, the model may be used to explore how seasonal and regional limitations of temperature, prey availability, and turbidity integrate to provide more or less favorable conditions for individual delta smelt.

Chipps and Wahl (2008) discuss the development and uses of bioenergetics models of fish populations. They recommend that management applications of bioenergetics models be limited by the degree to which model components have been validated. The IBMR was calibrated to certain metrics observed in the wild delta smelt population, relative abundance and mean individual growth rates, but other components of the model, discussed below, require empirical validation in a controlled (hatchery) environment. **Given critical uncertainties, the best use of the IBMR would be to qualitatively compare among simulated baseline and management action scenarios;** quantitative applications, such as predicting abundance, would not be an appropriate use of the IBMR at the current stage of model development and validation. Comparing simulated population growth rates (mean λ_{AB}) under a baseline scenario, represented by environmental, fish distribution, and prey observations in 1995-2014, to a management action scenario, represented by some modification to observed conditions, may illustrate long-term expected effects of management actions. The sensitivity analyses reported on here represented a version of this approach, in which mean λ_{AB} were evaluated when observed conditions were fixed at 1997 or 2005 values.

Bioenergetics models provide a method to integrate information about the attributes of the abiotic and biotic environment with mechanistic models describing the fitness consequence of using that habitat space (Rosenfeld et al. 2016). At the smallest spatiotemporal scale, comparisons of potential seasonal or monthly growth may help to illustrate regional differences abiotic and biotic conditions that lead to regional variation in habitat quality. Focusing on just bioenergetics, and avoiding the reproduction, mortality, and movement assumptions of IBMR, regional estimates of temperature and turbidity can be used to evaluate habitat suitability in terms of individual growth potential. This regional habitat suitability approach can be expanded to include biotic conditions if prey availability can be characterized. Comparing regional estimates of potential consumption, given measurements of abiotic and biotic conditions, may reveal that some areas provide greater productivity than others. The value of this approach is that it allows integration of several components of habitat quality into a single index of habitat suitability.

Model limitations and potential future improvements

The model primarily relies on the validity of experimental results to parameterize vital rates, such as fecundity and larval survival, and measurements of environmental quantities across the spatial domain of the delta smelt population. Future experimental results or the development of new datasets may lead to model improvements.

Though IBMR was not statistically fit to observed abundance data, it was compared to abundance estimates from the Delta Smelt Life Cycle Model with Entrainment (Smith et al. 2021), and this comparison illustrated certain aspects of model performance. Primarily, lack of fit to some June recruitment estimates suggested that IBMR generated greater variation in recruitment than was observed. This was related to the findings that IBMR was sensitive to prey density estimates, that IBMR-simulated female size in February was greater than observed in some years and less than observed in other years, and the finding of Rose et al. (2013b) that DSIBM was sensitive to fecundity estimates. Higher prey densities generate larger, more fecund simulated females, resulting in higher simulated recruitment in June. Improved quantification of bioenergetics model parameters for delta smelt will lead to improved representation of delta smelt growth and fecundity.

Estimation of full life cycle occupancy models would enable model-based simulations of fish distributions, which may change under new conditions that represent management actions. The IBMR v3 is restricted to observed fish distributions, and this approach ignores incomplete detection by the fish surveys. The ability to quantify the observed spatial distributions of delta smelt was limited by the low density of the population, which resulted in increasingly zero-inflated data near the end of the time series. 2014 was the terminal IBMR year, partially because delta smelt were undetected in most spatial strata by 2015. Development of occupancy models could also allow the extension of the simulation model to years when adequate survey data were not available, prior to 1995 or after spring of 2015.

Strata volumes, developed using DSM2, were used to expand observed densities to regional abundances, which were then used to parameterize the movement model. Total strata volumes may not represent delta smelt habitat, with strata having large volumes of deep water being over-weighted. Volume to four meters has been used in other applications to characterize delta smelt habitat (Polansky et al. 2019); however, volume estimates to four meters were not available for the strata delineations used in the IBMR.

Water temperature was estimated from data collected during fish monitoring in the Delta for years 2010-2014, and DSM2 was used to estimate temperatures in all years prior to 2010. DSM2 should be a better representation, because it includes a high level of spatial and temporal replication, whereas fish monitoring samples are collected a small number of times per year-month combination. Future model improvements should include DSM2 estimates of temperature for the entire IBMR time series, through year 2014.

The model assumed by Rose et al. (2013a) resulted in declining foraging rate as temperature increased above 23°C. Although the bioenergetics component of the model was extremely sensitive to the relationship between temperature and maximum foraging rate, or $C_{\max}$, the relationship has not been quantified empirically. An unpublished pilot study by Eder et al. (2014) suggested that delta smelt foraging rates may begin to decline at temperatures as low as 20°C, which would result in a greater thermal impact on bioenergetics than assumed in the IBMR.

The mechanistic representation of the population was limited by the monthly temporal scale. For example, temperatures greater than 23°C rarely occurred in the monthly IBMR $Temp_{\mathrm{yms}}$ dataset. At the daily or sub-daily scale, however, temperatures greater than 23°C occur more frequently, suggesting that extreme temperature effects were not well represented in the IBMR. The ability to simulate a daily time step is a major advantage of the DSIBM (Rose et al. 2013a and 2013b; Rose et al. 2021).

Turbidity effects were represented by greater simulated mortality and lower delta smelt consumption rates ($C_{\max}$) in clear water, indexed by Secchi depths greater than 84 cm. Experimental results suggest greater effects of turbidity on both mortality (Ferrari et al. 2014) and $C_{\max}$ (Baskerville-Bridges et al. 2004; Hasenbein et al. 2016), but we assumed these experimental results were biased by hatchery conditions and wild delta smelt would be more effective at avoiding predation and foraging in clear water. The maximum effect of turbidity, represented by parameter α_{FL}, was therefore calibrated in IBMR v3 to approximate the adult lengths observed in the wild population. Hasenbein et al. (2016) estimated a 60% reduction in foraging rates at low turbidity, and Ferrari et al. (2014) found a 66% lower survival rate at lower turbidities. In IBMR α_{FL} was calibrated to a value of 0.63. While experimental results guided our models of turbidity effects on delta smelt, the relationships effectively remain unquantified, representing a critical uncertainty in delta smelt bioenergetics models.

South Delta turbidity effects on the probability of delta smelt entrainment were documented by the US Fish and Wildlife Service's Life Cycle Model (LCM) (Smith et al. 2021). The prevailing hypothesis (Grimaldo et al. 2009) is that turbidity induces delta smelt to occupy portions of the water column where they become vulnerable to advective flows into the South Delta (OMR) if they happen to occupy or move into an area near the Lower San Joaquin River. Although LCM-based estimates of proportional entrainment were used in IBMR, the OMR and South Delta Secchi depth parameters quantified by the LCM could not be applied in IBMR, because the LCM was non-spatial and parameters described population-level effects. IBMR, on the other hand, was a spatial model requiring a spatially stratified movement model that cannot be derived from LCM parameters. IBMR entrainment was therefore modeled as a function of South Delta occupancy and OMR, ignoring the effect of South Delta turbidity in causing delta smelt to occupy portions of the Delta where entrainment was more likely. Failure to capture the effect of South Delta turbidity may be less important in simulating the effects of future management actions, if future conditions

are expected to resemble recent years, when entrainment has been lower and turbidity has been consistently low in the South Delta, especially during the spring.

Sensitivity analyses demonstrated that IBMR was sensitive to the abiotic and biotic conditions forming the simulated delta smelt foraging arena, prey density, temperature, and turbidity. Better measurements of these abiotic and biotic conditions, or better models to recreate past prey, temperature, and turbidity fields, will reduce once source of error in IBMR. For example, turbidity was indexed by Secchi depth, measured during fish monitoring. Low sample sizes within most year-month-strata combinations suggest that monthly measured Secchi depths were not always representative of ambient conditions. A method to recreate the past turbidity fields in the Delta, that were not consistently measured at the time, would improve one important representation of the delta smelt foraging arena.

Acknowledgements

The findings and conclusions in this report are those of the author and do not necessarily represent the views of the US Fish and Wildlife Service, the Department of Interior, or the other member agencies of the Interagency Ecological Program for the San Francisco Estuary (IEP). Li-Ming He, Matt Nobriga (USFWS), Sally Rudd, Brian Crawford, and two anonymous reviewers kindly provided comments on early drafts of this report.

References

Andrews, S. W., Gross, E. S., and Hutton, P. H. 2016. A water balance model to estimate flow through the Old and Middle River corridor. San Francisco Estuary and Watershed Science 14(2).

Bennett, W. A. 2005. Critical assessment of the delta smelt population in the San Francisco Estuary, California. San Francisco Estuary and Watershed Science 3(2).

Chipps, S. R., and D. H. Wahl. 2008. Bioenergetics modeling in the 21st century: reviewing new insights and revisiting old constraints. Transactions of the American Fisheries Society 137(1): 298–313.

Damon, L. J., S. B. Slater, R. D. Baxter, and R. W. Fujimura. 2016. Fecundity and reproductive potential of wild female Delta Smelt in the upper San Francisco Estuary, California. California Fish Game 102: 188–210.

Eder, K. J., R. C. Kaufman, D. E. Cocherell, J. C. Lindberg, N. A. Fangue, and F. J. Loge. 2014. Longfin and delta smelt food consumption and bioenergetics assessments. Final Report to U.S. Bureau of Reclamation Project # R10AC20107.

Ferrari, M. C., L. Ranåker, K. L. Weinersmith, M. J. Young, A. Sih, and J. L. Conrad. 2014. Effects of turbidity and an invasive waterweed on predation by introduced largemouth bass. Environmental biology of fishes 97(1): 79–90.

Haddon, M. 2001. Modelling and Quantitative Methods in Fisheries. Chapman and Hall/CRC, Boca Raton, FL.

Hasenbein, M., N. A. Fangue, J. Geist, L. M. Komoroske, J. Truong, R. McPherson, and R. E. Connon. 2016. Assessments at multiple levels of biological organization allow for an integrative determination of physiological tolerances to turbidity in an endangered fish species. Conservation physiology 4(1): 1–16.

He, L., K. Rose, D. Hilts, W. Kimmerer, and M. Nobriga. 2020. Updating and Expanding the Delta Smelt Individual-Based Life Cycle Model. Poster for the 2020 IEP Workshop.

Kimmerer, W., S. R. Avent, S. M. Bollens, F. Feyrer, L. F. Grimaldo, P. B. Moyle, M. Nobriga, and T. Visintainer. 2005. Variability in length–weight relationships used to estimate biomass of estuarine fish from survey data. Transactions of the American Fisheries Society 134: 481–495.

Kurobe, T., M. O. Park, A. Javidmehr, F. C. Teh, S. C. Acuña, C. J. Corbin, A. J. Conley, W. A. Bennett, and S. J. Teh. 2016. Assessing oocyte development and maturation in the threatened Delta Smelt, Hypomesus transpacificus. Environmental biology of fishes 99(4): 423–432.

LaCava, M., K. Fisch, M. Nagel, J. C. Lindberg, B. May, and A. J. Finger. 2015. Spawning behavior of cultured Delta Smelt in a conservation hatchery. North American Journal of Aquaculture 77(3): 255–266.

Nobriga, M. L., and W. E. Smith. 2020. Did a shifting ecological baseline mask the predatory effect of striped bass on delta smelt? San Francisco Estuary and Watershed Science

Pangle, K. L., T. D. Malinich, D. B. Bunnell, D. R. DeVries, and S. A. Ludsin. 2012. Context-dependent planktivory: interacting effects of turbidity and predation risk on adaptive foraging. Ecosphere 3(12):114.

Peterson, J., E. McCreless, A. Duarte, S. Hamilton, and J. Medellin-Azuara. 2019. Structured Decision Making for Scientific Management in the San Francisco Bay-Delta. Draft progress report, US Bureau of Reclamation, Sacramento, CA.

Polansky, L., L. Mitchell, and K. B. Newman. 2019. Using multistage design-based methods to construct abundance indices and uncertainty measures for delta smelt. Transactions of the American Fisheries Society 148(4): 710–724.

Polansky, L., Newman, K. B., and Mitchell, L. 2020. Improving inference for nonlinear state-space models of animal population dynamics given biased sequential life stage data. Biometrics.

R Core Team. 2018. R version 3.5.0: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rose, K., W. Kimmerer, D. Hilts, L. He, and M. Nobriga. 2021. An Individual-Based Model of Delta Smelt Population Dynamics: A Versatile Tool for Life-Cycle Analyses of Management Actions. Oral presentation for the 2021 Bay-Delta Science Conference.

Rose, K. A., W. J. Kimmerer, K. P. Edwards, and W. A. Bennett. 2013a. Individual-based modeling of delta smelt population dynamics in the upper San Francisco Estuary: I. Model description and baseline results. Transactions of the American Fisheries Society 142: 1238–1259.

Rose, K. A., W. J. Kimmerer, K. P. Edwards, and W. A. Bennett. 2013b. Individual-based modeling of Delta Smelt population dynamics in the upper San Francisco Estuary: II. Alternative baselines and good versus bad years. Transactions of the American Fisheries Society 142: 1260–1272.

Rosenfeld, J., H. Beecher, and R. Ptolemy. 2016. Developing bioenergetic-based habitat suitability curves for instream flow models. North American Journal of Fisheries Management 36(5): 1205–1219.

Smith, W. E. 2018. Delta Smelt Model Technical Note 35: Delta smelt growth model. US Fish and Wildlife Service, Bay-Delta Office, Sacramento, CA. Technical Report.

Smith, W. E, L. Polansky, and M. L. Nobriga. 2021. Disentangling risks to an endangered fish: using a state-space life cycle model to separate natural mortality from anthropogenic losses. Canadian Journal of Fisheries and Aquatic Sciences. *In press*. doi: 10.1139/cjfas-2020-0251.

Sommer, T., C. Armor, R. Baxter, R. Breuer, L. Brown, M. Chotkowski, S. Culberson, F. Feyrer, M. Gingras, B. Herbold, and W. Kimmerer. 2007. The collapse of pelagic fishes in the upper San Francisco Estuary: El colapso de los peces pelagicos en la cabecera del Estuario San Francisco. Fisheries 32(6): 270–277.

Table 1. Comparison of individual-based life cycle model (LCM) variations for delta smelt. Each of the models shares some of the population dynamics functions and the same embedded bioenergetics model and parameters. Each has unique features as well, particularly their details of how movement, entrainment, and reproduction are modeled.

Model Version (citation)	Programming Language	Details
DSIBM (Rose et al. 2013)	FORTRAN90	Full LCM on daily time step simulates 11 overlapping generations in 11 spatial strata; vital rate outcomes are predominantly mechanistic or calibrated; fish movement is the emergent interaction of behavior rules and hydrodynamics.
IBMR v1 (Smith 2018)	R	Partial LCM (no reproduction) on monthly time step simulates 11 non-overlapping generations in 11 spatial strata; vital rate outcomes are mechanistic but uncalibrated; fish movement is based on spatial distributions observed in historical catch data.
IBMR v2 (Peterson et al. 2019)	R	Full LCM on monthly time step simulates 20 non-overlapping generations in 12 spatial strata; vital rate outcomes are mechanistic and calibrated; fish movement is based on occupancy modeling of historical catch data.
DSIBM v2 (Rose et al. 2021)	FORTRAN90	Full LCM on daily time step simulates 21-82 overlapping generations in 12 spatial strata; vital rate outcomes are predominantly mechanistic or calibrated; fish movement is the emergent interaction of behavior rules and hydrodynamics.
IBMR v3 (Smith 2021; this document)	R	Full LCM on monthly time step simulates 20 non-overlapping generations in 12 spatial strata; vital rate outcomes are mechanistic and calibrated to entrainment mortality, abundance, and growth rates observed in wild delta smelt; fish movement is based on spatial distributions observed in historical catch data; updated with recently developed life history parameters.

Table 2. All indices, data, and simulated parameters used to denote IBMR.

Index	Description	Values
y	year	y = 1995, 1996, …2014
m	month	m = 2, 3, …12, 1
s	spatial strata	s = 1, 2, …12
p	prey type	p = 1, 2, …12
i	individual	i = 1, 2, …
Data	Description	
PD_{ymsp}	prey density	
OMR_{ym}	Old and Middle River flow	
DS_{yms}	observed delta smelt spatial distribution	
$Temp_{\mathrm{yms}}$	water temperature	
$Secchi_{\mathrm{yms}}$	Secchi depth	
$R_{\mathrm{ss'}}$	matrix of movement rules between strata *s* and *s'*	
Simulated variables	Description	
L_{iym}	fork length	
W_{iym}	weight	
$mature_{\mathrm{iym}}$	maturity status	
$female_{\mathrm{i}}$	sex	
$n_{\mathrm{AD_{yms}}}$	abundance of adults	
$n_{\mathrm{Egg_{yms}}}$	number of eggs produced	
$S_{\mathrm{LRV_{ym}}}$	survival of eggs to the larval life stage	
$P_{\mathrm{Hatch_s}}$	probability of egg hatching	
$S_{\mathrm{YS_{yms}}}$	survival of yolk-sac larvae	
$L_{\mathrm{LRV_i}}$	larval length at first feeding	
$strata_{\mathrm{iym}}$	stratum occupied by individual *i*	
C_{iym}	realized consumption rate	
R_{iym}	respiration	
F_{iym}	egestion	
U_{iym}	excretion	
SDA_{iym}	activity	
Sp_{iym}	spawning loss	
C_{max}	maximum potential consumption	
$S_{\mathrm{JA_{iym}}}$	survival of juveniles through adults	
F_{iym}	instantaneous rate of entrainment mortality	
M_{iym}	instantaneous rate of natural mortality	
z_{iym}	survival status	
F_{max}	maximum potential rate of entrainment mortality	
z_{iym}	survival status	
D_{yms}	number of simulated fish per m^3 (density)	
$\lambda_{\mathrm{AB_y}}$	population growth rate	

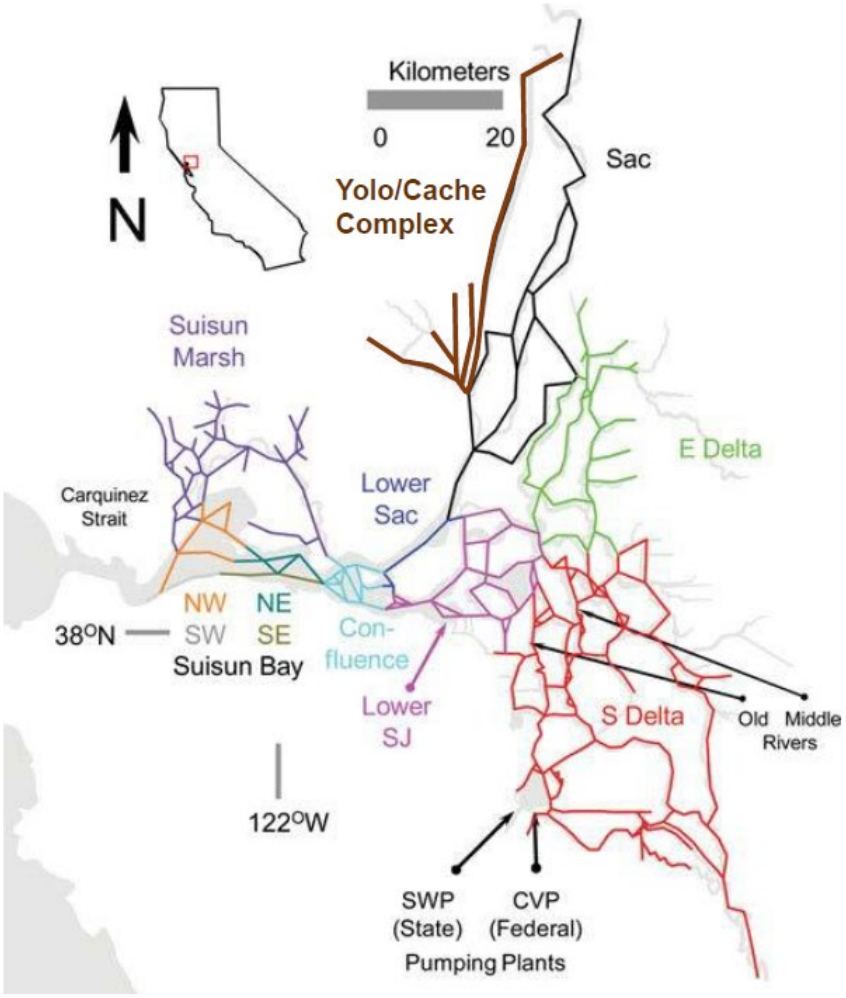


Figure 1. Map of the Sacramento-San Joaquin Delta, showing the spatial strata used to model delta smelt spatial distributions. This map was reproduced from Rose et al. (2013a) and Peterson et al. (2019).

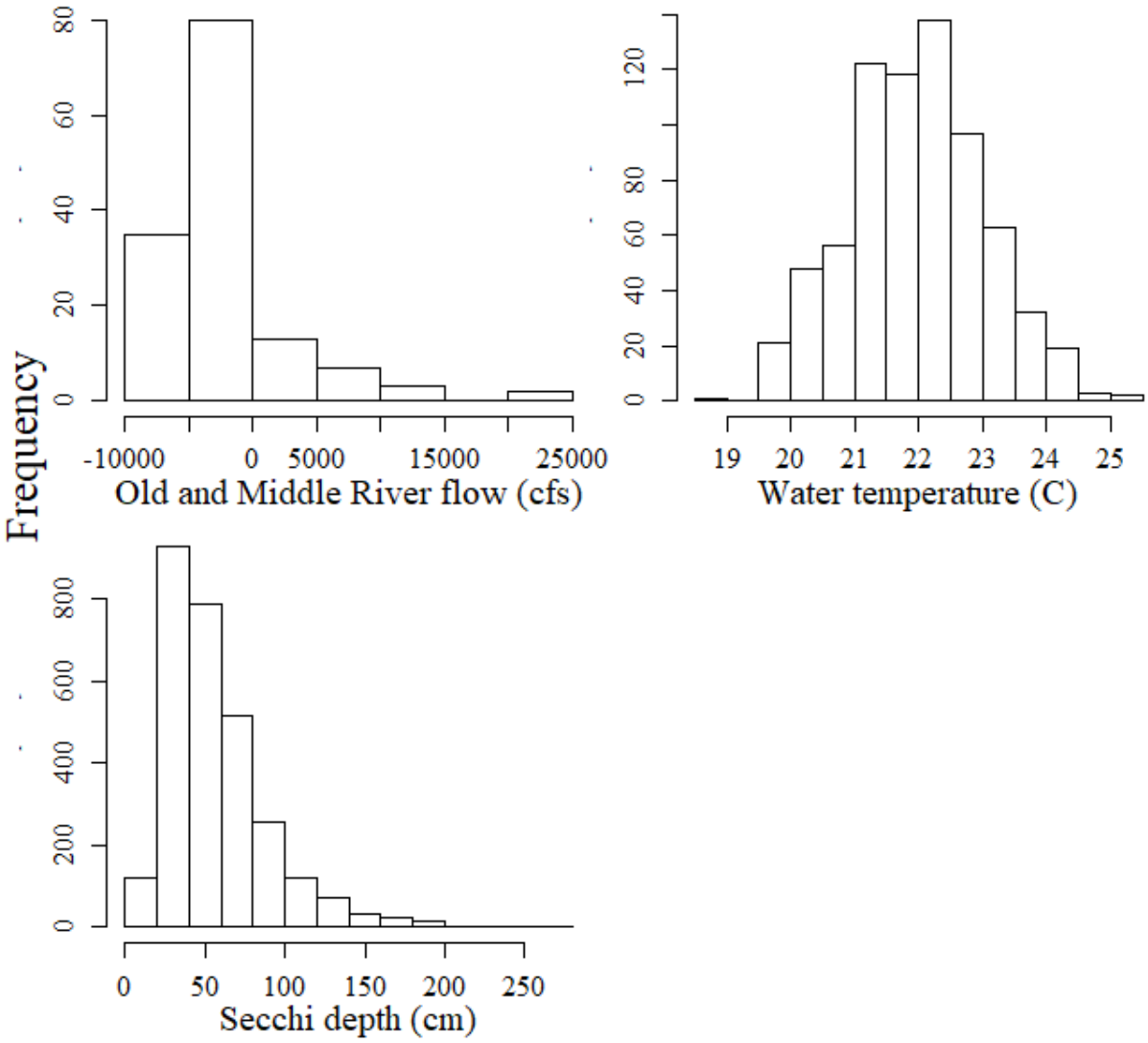


Figure 2. Histograms of physical Delta variables driving the Delta Smelt Individual-Based Life Cycls Model in R (IBMR). Old and Middle River flows for the critical December through June period and water temperatures for July through September are shown, because these variables were expected to be seasonally limiting. Secchi depths for all year-month-stratum combinations are shown.

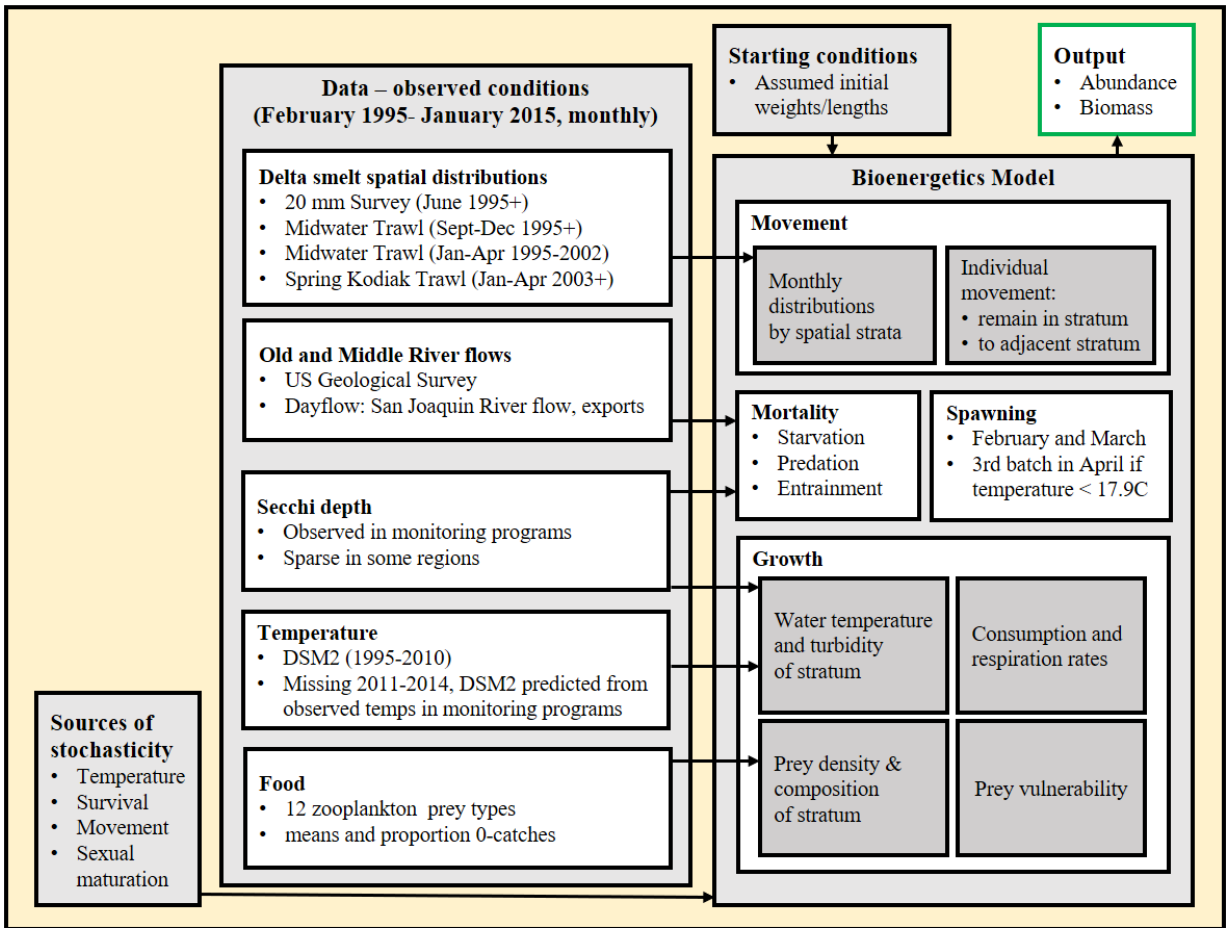


Figure 3. Diagram of the Delta Smelt Individual-Based Life Cycle Model in R (IBMR) and data.

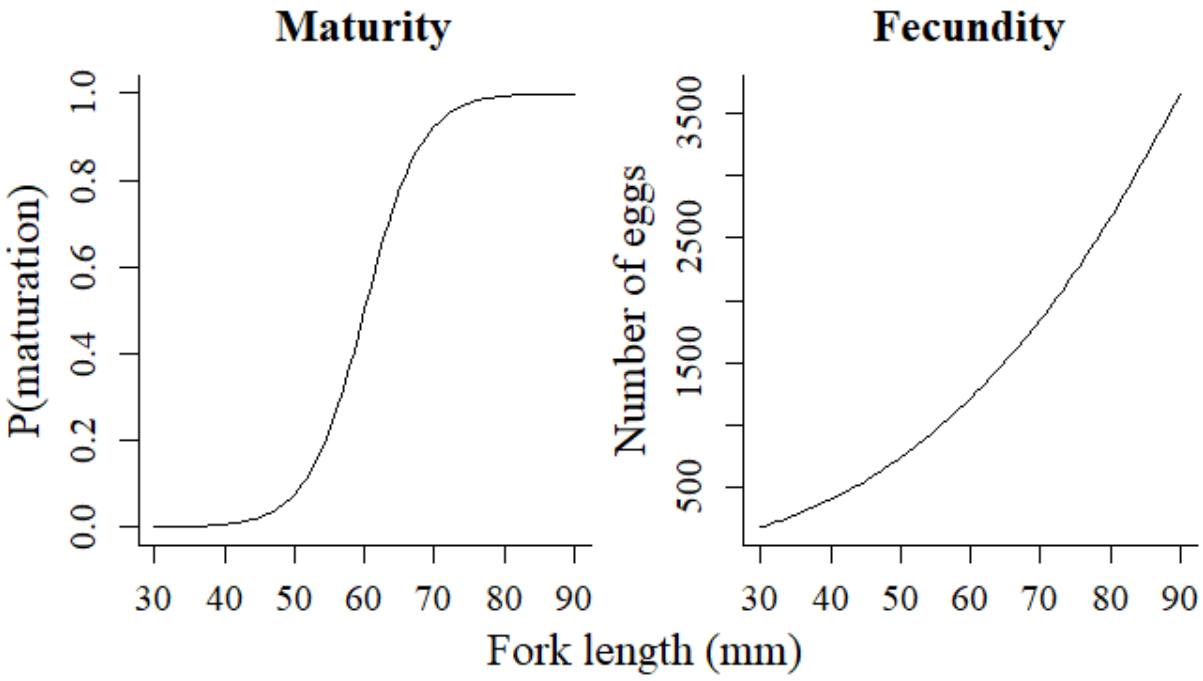


Figure 4. Models of the probability of maturation and fecundity at length.

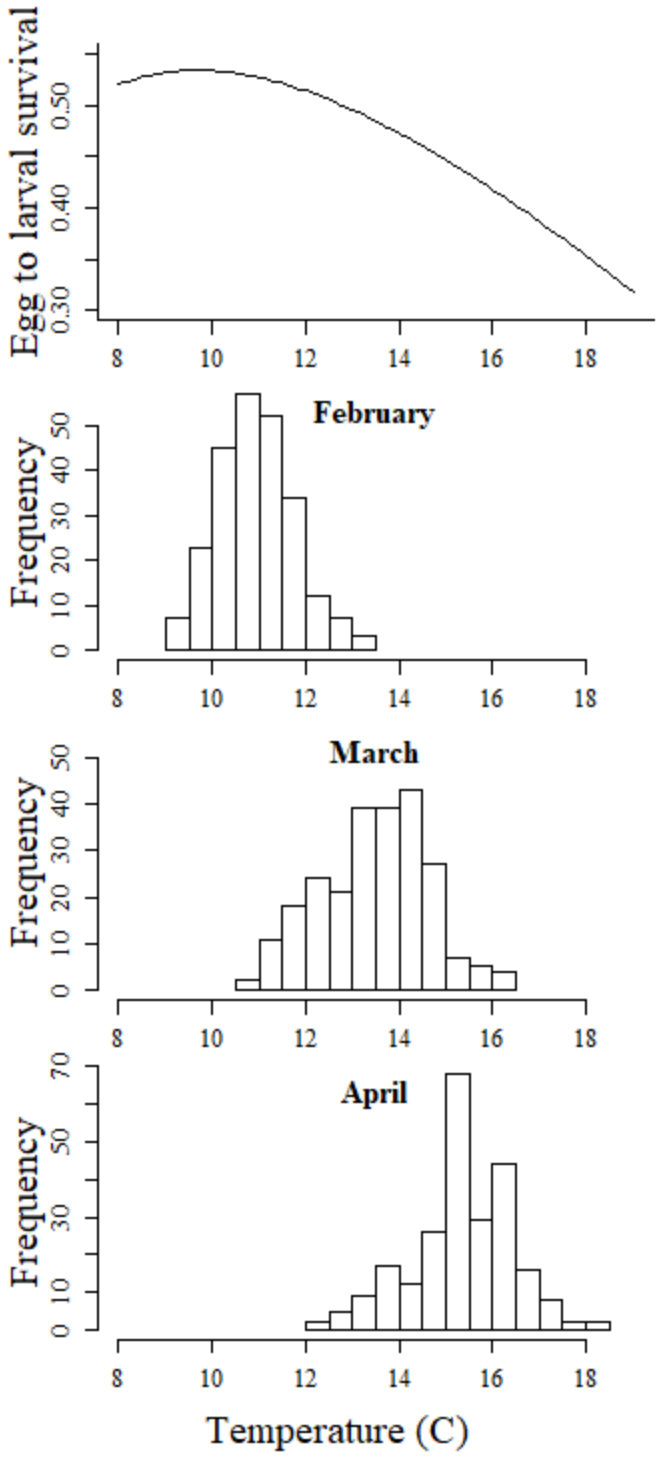


Figure 5. Model of egg to larval survival as a function of water temperature (top panel), based on Bennett (2005) (Bennett's Fig. 10). The bottom three panels show observed Delta temperatures in all spatial strata during spawning months of 1995-2014.

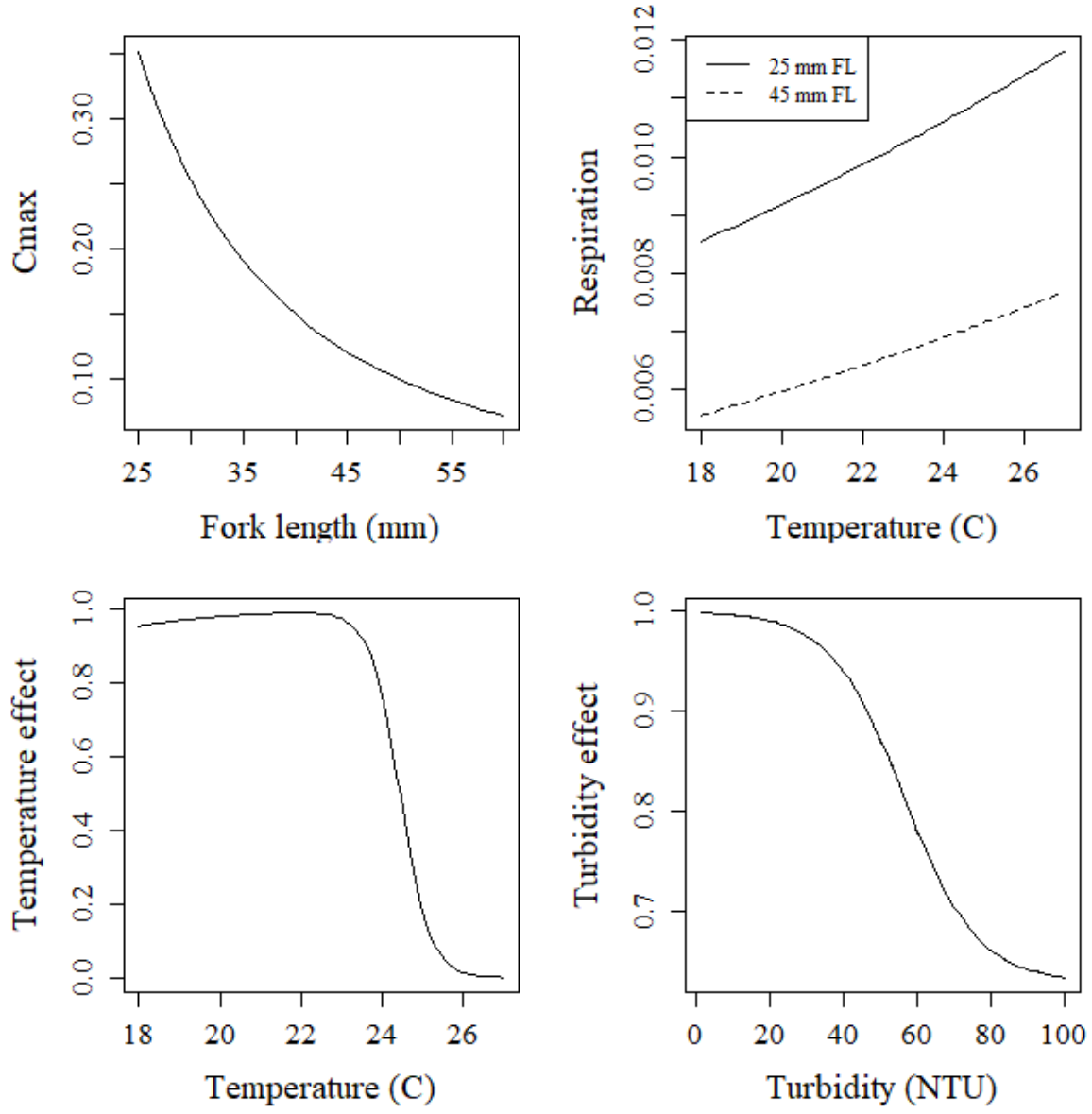


Figure 6. Models of maximum consumption (C_{max}) and respiration assumed by Rose et al. (2013a) (top row). In the bottom row are shown models of the temperature effect on C_{max} ($\boldsymbol{KA}_{\mathbf{iym}}$ and $\boldsymbol{KB}_{\mathbf{iym}}$; Eq. 12-13), and the model of the effect of turbidity on C_{max} ($\boldsymbol{KS}_{\mathbf{iym}}$; Eq. 14), suggested by data published by Hasenbein et al (2016).

TABLE 1. Parameter values for each Delta Smelt life stage in the bioenergetics model.

Parameter	Description	Larvae	Postlarvae	Juveniles and adults
	Maximum consumption (C_{max})			
a_c	Weight multiplier	0.18	0.18	0.1
b_c	Weight exponent	−0.275	−0.275	−0.54
CQ (°C)	Temperature at CK_1 of maximum	7	10	10
T_O (°C)	Temperature at 0.98 of maximum	17	20	20
T_M (°C)	Temperature at 0.98 of maximum	20	23	23
T_L (°C)	Temperature at CK_4 of maximum	28	27	27
CK_1	Effect at temperature CQ	0.4	0.4	0.4
CK_4	Effect at temperature T_L	0.01	0.01	0.01
	Metabolism (R)			
a_r	Weight multiplier	0.0027	0.0027	0.0027
b_r	Weight exponent	−0.216	−0.216	−0.216
R_Q	Exponent for temperature effect	0.036	0.036	0.036
S_d	Fraction of assimilated food lost to SDA	0.175	0.175	0.175
	Egestion (F) and excretion (U)			
F_a	Fraction of consumed food lost to egestion	0.16	0.16	0.16
U_a	Fraction of assimilated food lost to excretion	0.1	0.1	0.1

Figure 7. Table from Rose et al. (2013a) showing fixed parameter values used to simulate Delta Smelt feeding and growth.

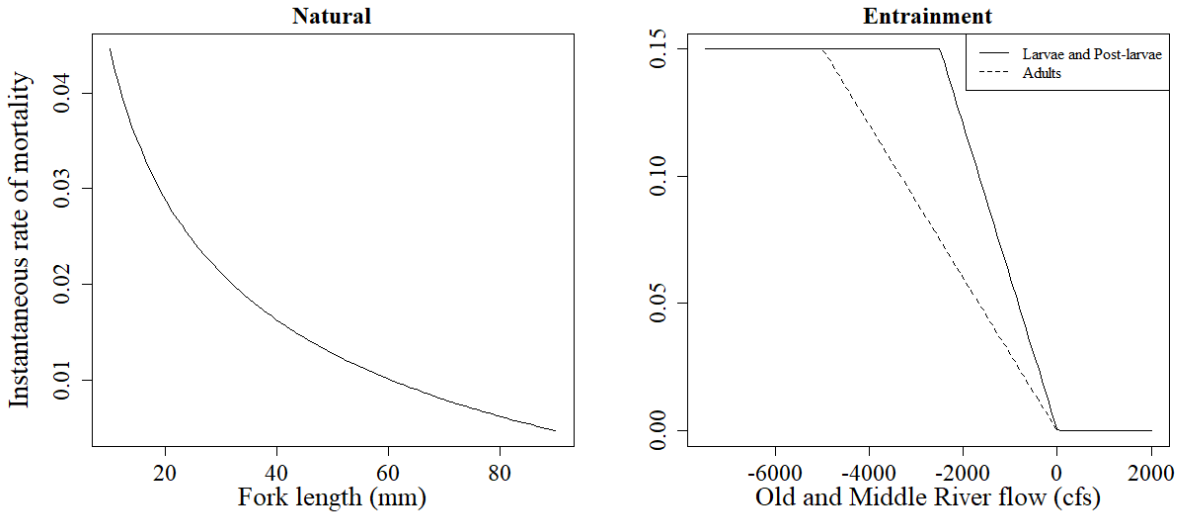


Figure 8. Models of instantaneous rates of natural (~ predation) and entrainment mortality. Entrainment mortality risk was added to natural mortality risk if fish occupied the South Delta stratum.

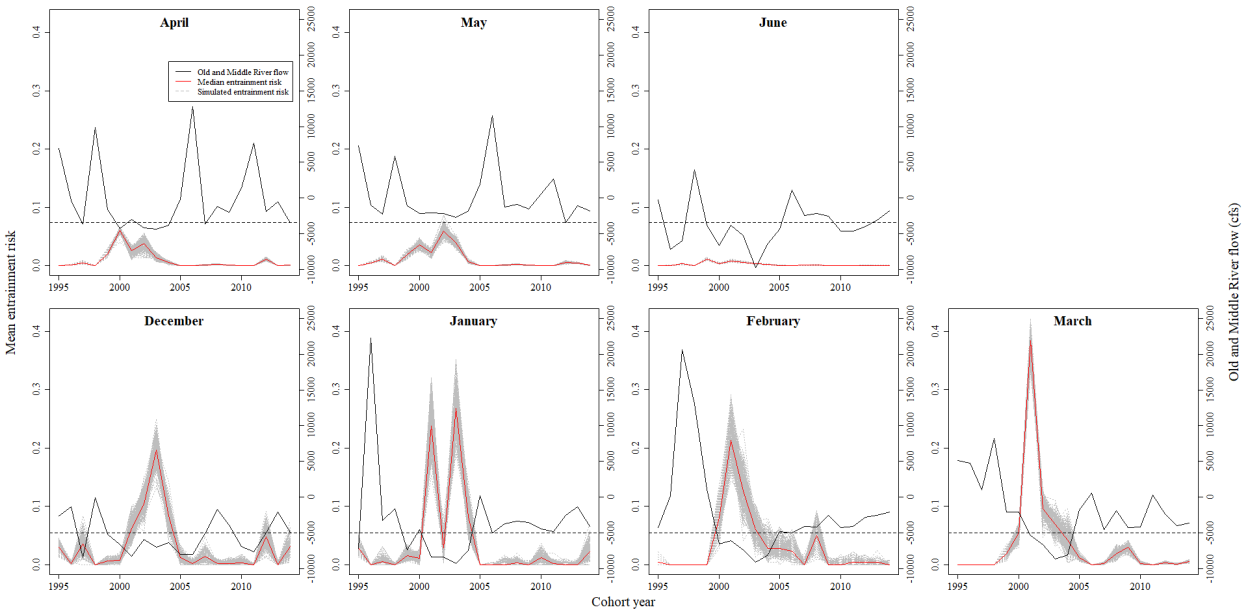


Figure 9. Median entrainment mortality (red lines), among all simulated populations (gray lines) during January through May of each year. Old and Middle River flow experienced by the simulated delta smelt population is shown by the black lines.

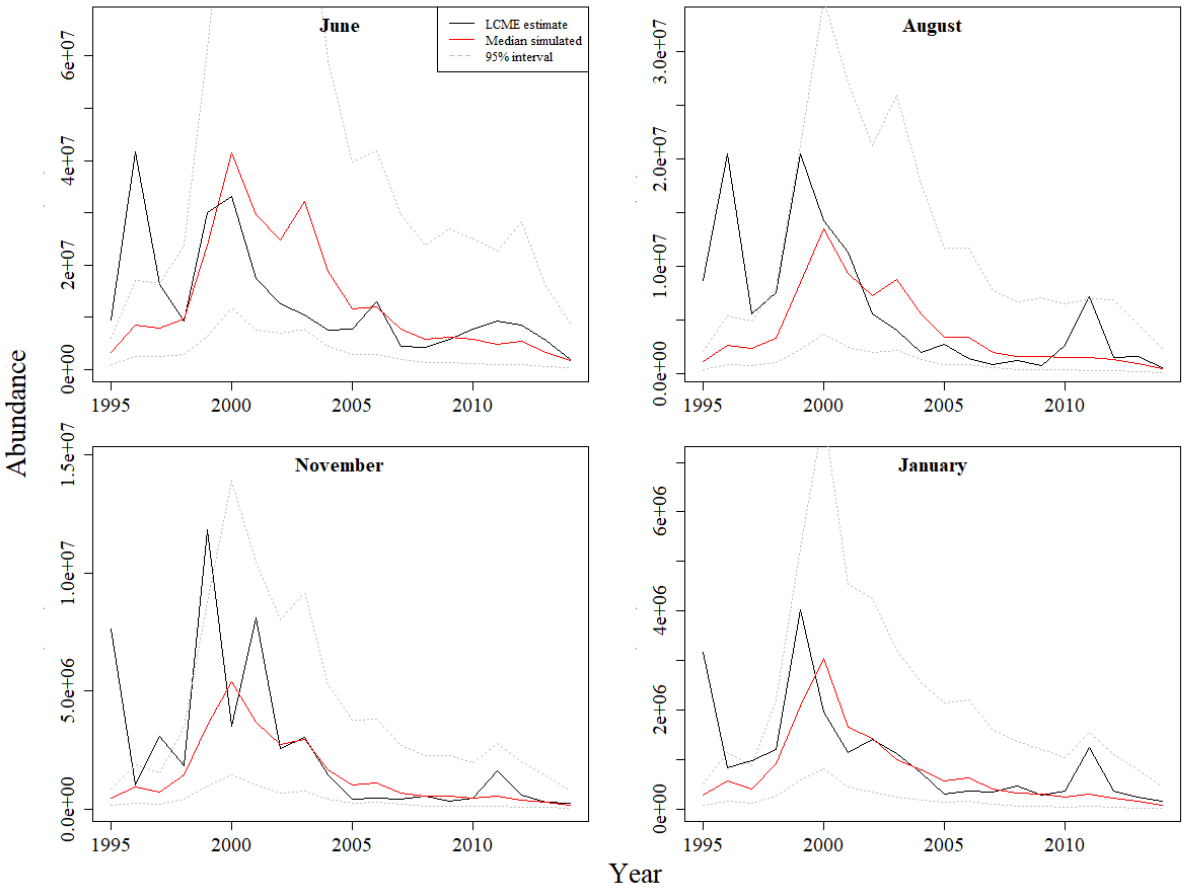


Figure 10. Median simulated abundance (red lines), among two simulated population time series (gray lines), and abundances estimated the Delta Smelt Life cycle Model with Entrainment (black lines) (Smith et al. 2021).

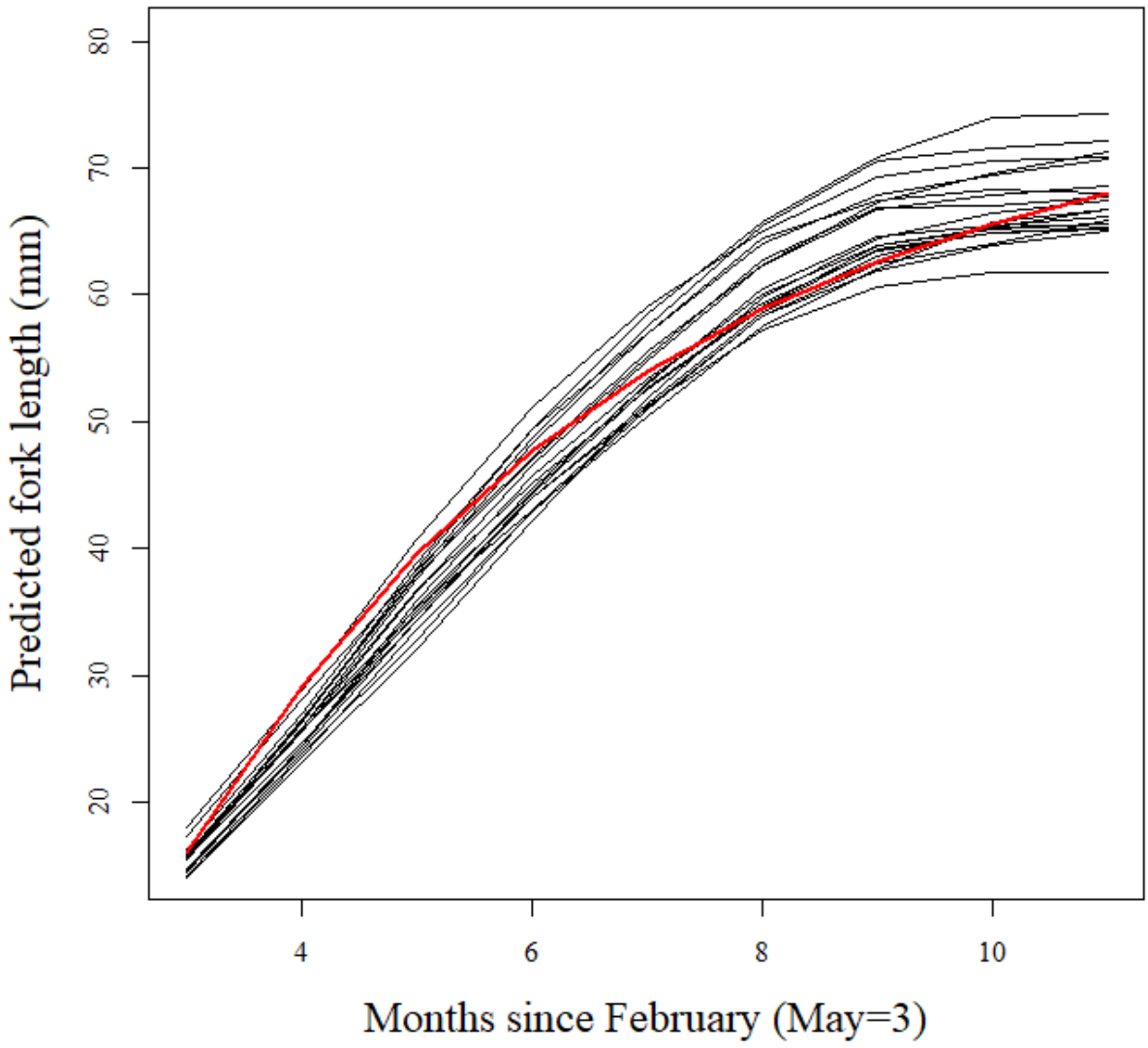


Figure 11. Simulated monthly time series of mean delta smelt length at age, for all years 1995-2015 (black lines) and von Bertalanffy growth model predicted length at age (red line), assuming a mean length of 16 mm FL in May. The von Bertalanffy model was fit to lengths and ages observed in the wild population.

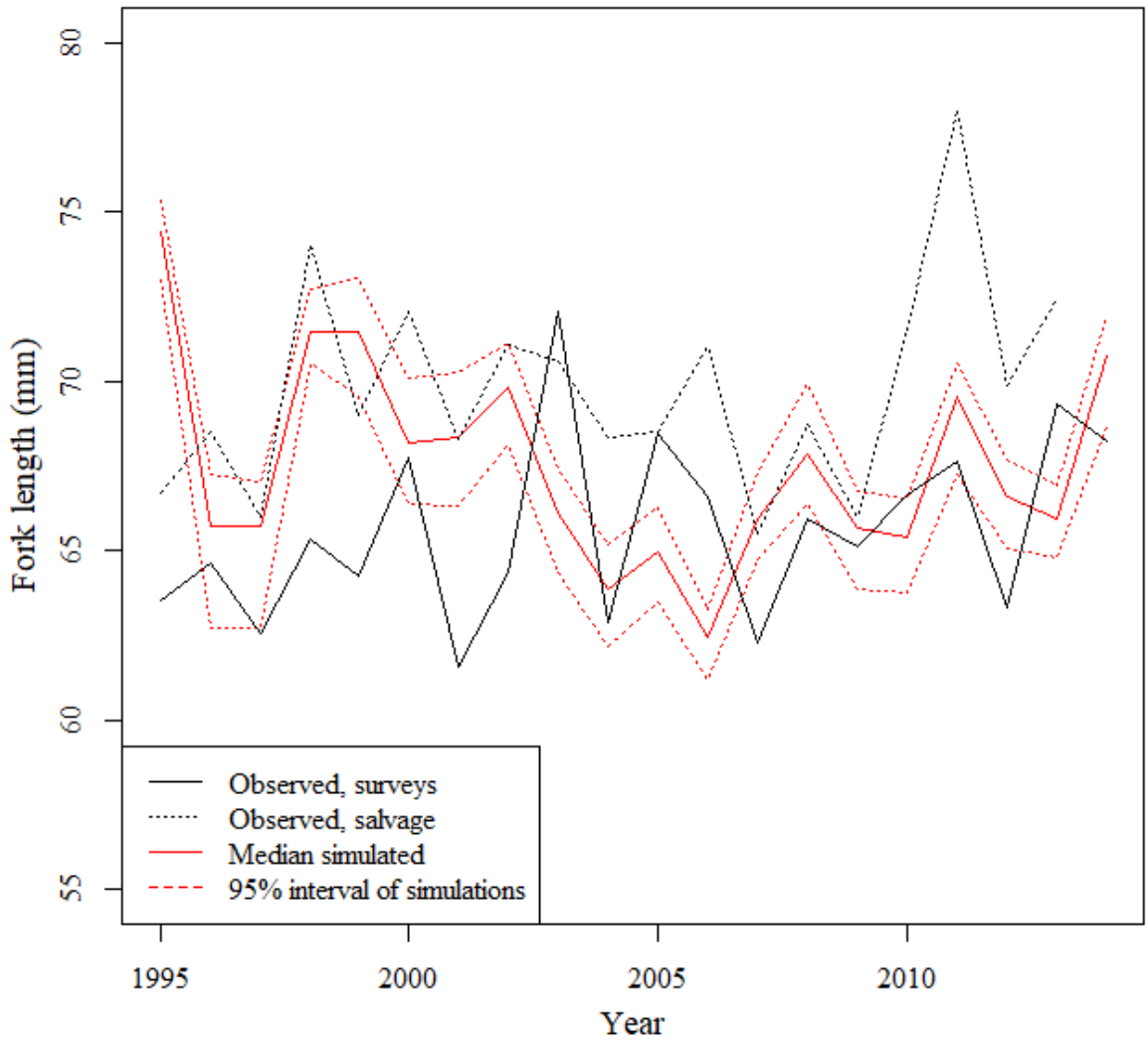


Figure 12. February fork lengths simulated by IBMR and observed for wild delta smelt in the CDFW Spring Kodiak (2002-2014), Midwater (1995-2001) Trawl Surveys, and in salvage operations.

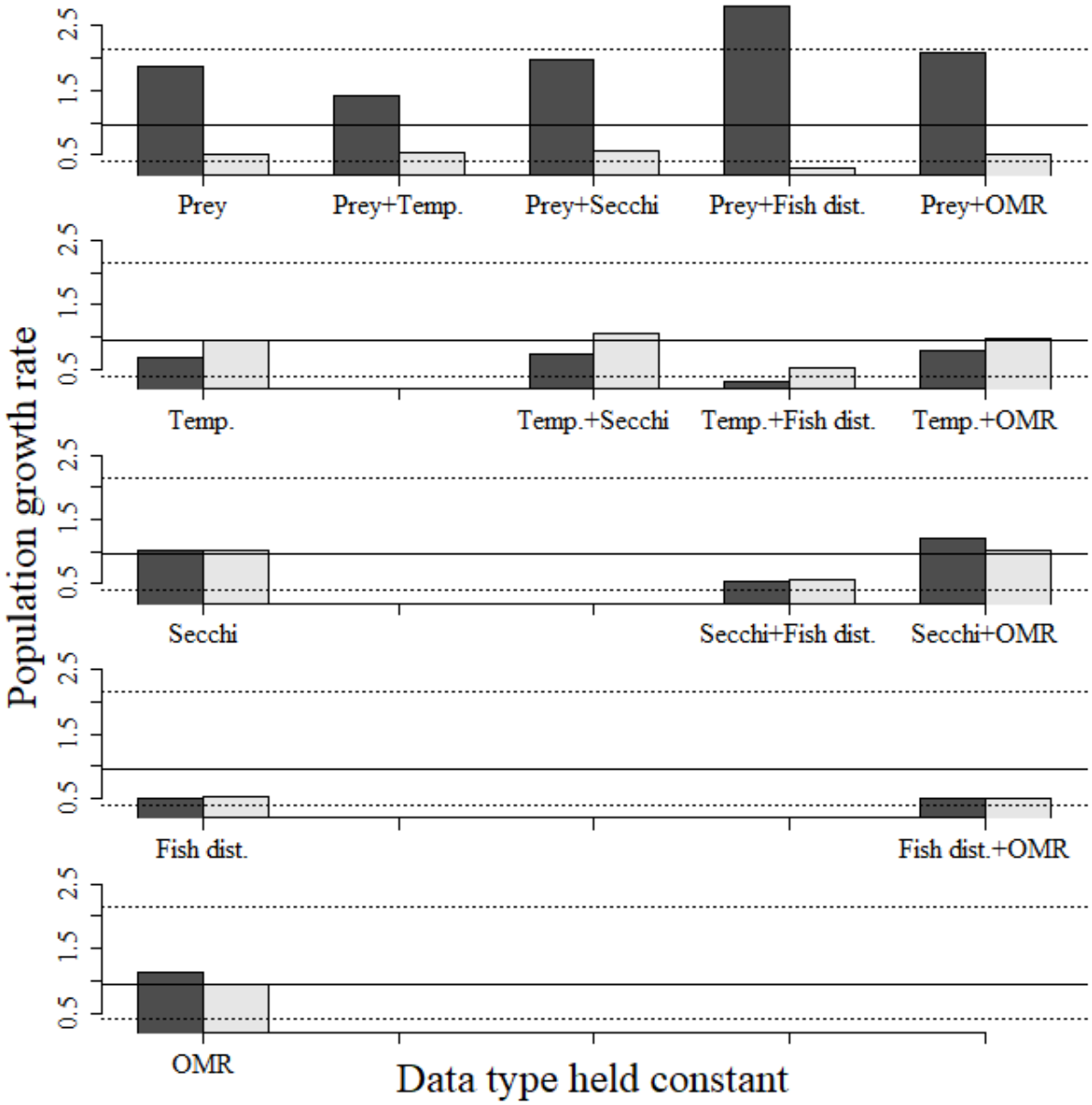


Figure 13. Results of sensitivity analysis. Bars show the geometric mean of population growth rates $\boldsymbol{\lambda_{AB}}$, when one or two data types were fixed at the values measured in high growth year 1998 (dark bars) or low growth year 2004 (light bars). Data types were prey density, temperature (Temp.), Secchi depth, observed delta smelt spatial distribution (Fish dist.), and Old and Middle River flow (OMR). Horizontal reference lines indicate the mean $\boldsymbol{\lambda_{AB}}$ when all data was allowed to vary at the values measured in 1995-2014 (solid line) and the maximum and minimum $\boldsymbol{\lambda_{AB}}$ simulated in 1998 and 2004 when all data was allowed to vary (dotted lines).

Appendix A

Methods for preparing zooplankton data for model input

Wim Kimmerer and Kenny Rose
27 December 2019

The methods used to populate the previous version of the model with inputs of zooplankton food for delta smelt (Appendix A in Rose et al. 2013a) have been altered in various ways to accommodate the objectives and design of this project (Table 1). In the previous version, zooplankton in six taxonomic groups were used to represent food of delta smelt. Data on abundance (number m^{-3}) were obtained from the Interagency Ecological Program zooplankton monitoring database and the CDFW 20-mm survey. The zooplankton program has collected zooplankton abundance data monthly at 16–22 stations during the period represented by the model of 1995–2005. The 20-mm survey has sampled at 49–52 stations each year, generally during March or April to July. These were used with estimates of carbon mass per individual to obtain estimates of biomass (mgC m^{-3}) for each taxon (species or higher taxonomic level) and major life stage. Data for each taxon, spatial box, and sampling date were increased by a small number (to allow for zeros), log-transformed, and averaged. Then data were interpolated to each date in the model run by the use of a moving window of 45 days on each side of the date and averaging all values within that window for the taxon and box. A similar process was used to determine standard deviations.

Our expansion of the spatial and temporal frame of the model, together with new information about delta smelt and their habitat, led us to reconsider the sources of data and the method of converting the available data to model input.

Objective: to extend the time frame of the zooplankton data to include water years 1991 through 2011, incorporate all available zooplankton data, expand the zooplankton groups, improve the method of assigning values to boxes, and revise the parameters for prey consumption as a step in the calibration process.

Overview: Data on zooplankton abundance were obtained from all available data sets from ongoing monitoring, and converted to biomass using previously determined carbon masses per individual by species and life stage (Fig. 1). Taxa were assigned to one of 12 taxonomic and life stage prey groups ("taxa"), and stations were assigned to spatial boxes in the IBM. Biomass data were converted to matrices of proportion zero catch and log mean of non-zero data, for each day of the simulation, spatial box, and taxon for input to the IBM. Then the parameters of the functional response of the four major delta smelt life stages were determined iteratively to give reasonable proportions of prey taxa consumed and gave a mean overall consumption rate of 75% of maximum during the calibration period (Fig. 2). These parameters were also used as input to the IBM.

Data sources: Zooplankton data are available from four monitoring programs (Tables 1, 2). The three fish-monitoring programs take zooplankton samples mainly to provide information for studies of fish diets. The durations, seasons, and stations differ among these programs, and in particular the spatial extent of these programs differs widely. Sample collection and processing is

more consistent among the programs, with a slightly larger mesh size used in the three fish programs, and some slight differences in sample processing and levels of species identification. In the zooplankton monitoring program an additional pump sample is taken to collect the smaller fraction (45–154 µm), which is the only suitable source of data for two of the prey taxa, copepod nauplii and *Limnoithona* spp.

Taxon codes in the three fish programs are based on those in the longer-running zooplankton program, so for most taxa there is no ambiguity. However, the level of identification differs somewhat among programs and has changed through time in all programs, presenting a challenge for finding a common description of the zooplankton assemblage. Adult copepods are usually (not always) identified to species, while juveniles are usually identified to genus; for most genera there is only one abundant species in the model domain.

Data processing: Processing the data required two principal steps: assembling the data sets and reconciling differences in identification. The data are available from CDFW through an FTP site (accessible through http://wdl.water.ca.gov/iep/products/data.cfm). The zooplankton data are provided as Excel files for each of the sampling gears. The fish data and associated zooplankton data are in Microsoft Access files. For analysis I exported all tables and imported them into R, then saved the data as lists of data tables in .Rdata files and wrote query functions to extract the data of interest from these lists. This resulted in smaller files and easier queries than available in Access. Data from the summer townet and fall midwater trawl surveys have identical formats, so these were combined.

We reconciled taxon names among sampling programs by merging the taxon lists from all programs and then examining the resulting table for taxa that were missing from one or two of the three data sets. These were adjusted as follows: 1) If one name was used for different taxa among data sets, the name was changed in one or two data sets to match the other(s); 2) If the taxon was either too large (e.g., mysids) or too small (e.g., rotifers) to be collected quantitatively in the net samples, the taxon was eliminated; or 3) If the taxon was described at different levels of resolution in different data sets, the coarser level of resolution was used.

Small forms such as copepod nauplii and small copepods (e.g., *Limnoithona* sp.) are collected int the zooplankton program using a pump sampler. Although these are also reported by the fish-monitoring programs, the nets used do not collect them quantitatively so data for these taxa were taken only from the pump data taken by the zooplankton program. Data for large and small zooplankton were processed separately because of the large difference in sampling density.

Stations: The association between fixed stations and boxes was determined previously for Rose et al. (2013a), except that we removed the Cache Slough complex from Box 1 (upper Sacramento) and placed it in new Box 12. Stations were assigned to boxes they were in and also to nearby boxes to increase the number of data points for each sampling day (see Appendix in Rose et al. 2013a).

Starting in 1994 the zooplankton monitoring program has also visited movable stations, one at salinity ~1 and the other at salinity ~3.2. When these salinities were found in the Delta, samples were taken in the Sacramento River, except that beginning in 2014 samples were also taken at these salinities in the San Joaquin River. One or more of these stations are omitted if the salinity

at a fixed station (the substitute) is close to the target value. Beginning in 2004 the coordinates of the movable stations and any substitute station were listed in a series of spreadsheets. We eliminated movable stations that had a substitute (to avoid double-counting data) and used these coordinates to assign each movable station to the nearest box.

For data before 2004 we compared the salinity recorded for the movable stations with that recorded at fixed stations in the same survey, and eliminated four data points from movable stations where the salinity matched . Then for each survey we constructed a relationship of salinity vs. distance up the axis of the estuary (km) and interpolated to convert salinity recorded at the movable stations to distance. This was then used to assign the movable stations to boxes. In four cases salinity was not recorded at any station, and these four values were dropped.

Selecting taxonomic groups

Equation 10 in Rose et al. (2013a) specifies the realized ingestion rate of a given prey group by an individual fish, rearranged here:

$$C_j' = \frac{C_j}{WC_{max}} = \frac{\frac{Z_j V_j}{K_j}}{1 + \sum_{\text{all k}} \frac{Z_k V_k}{K_k}} \quad (1)$$

where C_j' is the ratio of ingestion rate to the maximum C_{max} scaled to the weight of the fish W; that is, C_j' is a measure of the "success" of the fish because the growth rate of the fish will be maximized when C_j' is 1. Z_j is biomass of taxon j, V_j is vulnerability, a switch (0 or 1) to turn feeding on the species group on or off, and K_j is the half-saturation constant. Subscripts j refer to the individual prey group, which may be one or more life stages, species, or other groupings, and k refers to all prey groups. With only one prey group Eqn. 2 reduces to the familiar Holling Type II functional response, a rectangular hyperbola, in which the maximum is 1 and the half-saturation constant is K_j.

If two or more species groups have the same V and K parameters, total mass consumption by a fish does not change if the groups are entered individually or combined. Therefore species groups need be entered separately only if their parameters differ for any life stage of delta smelt. Other groups may also be entered separately for convenience during calibration or when comparing among time periods (above), boxes, or sampling programs.

The criteria for selecting taxonomic groups and either lumping them or keeping them separate are:

1. Relatively abundant during at least part of the model period.
2. Abundant in all sampling programs or in one or more boxes.
3. Known or likely difference in feeding parameters for any delta smelt life stage
4. Different spatial and temporal distributions from other taxa that could affect delta smelt
5. No more than 12 prey groups

Based on these criteria we selected 12 zooplankton groups (Table 3). Copepods include adults of two species, adults and juveniles of two species combined (*Limnoithona* spp.), juveniles of three species (*Pseudodiaptomus* spp.), other calanoid adults and juveniles, other cyclopoids, and copepod nauplii (larvae). Additional taxa are *Daphnia* spp. cladocera, other cladocera, and other taxa.

Developing master zooplankton data set

The sampling programs report abundance (number m^{-3}) based on counts of subsamples. I converted these to biomass using carbon mass measured on several adult copepods and other taxa (e.g., Kimmerer 2006, Gould and Kimmerer 2010, Kimmerer et al. 2017). Copepodites (juvenile copepods) were assumed to have about 25% of the carbon mass of adults based on their median life stage (copepodite 3) and the ratios of copepodite to adult masses for several species discussed in the above references. Masses for several other taxa were estimated from their size using literature values.

Large zooplankton were collected with 154–160 µm mesh nets towed obliquely through the water column. I used the combined data from all zooplankton monitoring studies. The first steps were to remove taxa that were not reported in the EMP zooplankton database and those with a mean biomass among all samples less than 0.05 mgC m^{-3}. Then I examined the biomass of each taxon in the three databases, as well as biomass in the low-salinity zone and the Cache Slough complex. There was general agreement among the databases and locations for most taxa.

Biomass data were natural-log-transformed with zeros in the original data flagged and replaced with a value of -5 which is below the minimum non-zero value. Data for each box, taxon, and sampling date were placed in two arrays, the first containing the log-transformed data excluding zeros (or -5 if all were zero). The second array contained a value of 1 for each zero value in the raw data. Then data in each array were averaged across all samples within a box, taxon, and date, though not all dates had complete samples. In addition box 12 was sampled only beginning in 1995, and then only in the southern portion of the box and during spring. In 2005 additional sampling began in summer–fall.

To fill out the data for Box 12 I first examined the data for correlations among boxes. Biomass in boxes 2-5 had strong relationships with that in Box 12, depending on the taxon. We therefore calculated models in which Box 12 biomass was modeled as a linear combination of data from Boxes 2-5. These models were used to predict the mean of the log-transformed non-zero values for Box 12 when data were missing. The proportion of zeros could not be modeled this way because much of the available data in Box 12 comprised only a single sample, but all the taxa had strong seasonal patterns of presence/absence. Therefore for each taxon we modeled the proportion of zeros as a smoothed function of julian day with a binomial error distribution (function *gam* in R package *mgcv*) and used that to predict the proportions of zeros for each taxon.

The next step was to extend the data for each sampling event in each box into each day of the model period. As before (Rose et al. 2013a) we applied a moving window of ± 45 d around each date, and calculated means of all non-missing data for a given box and taxon within that window. Because sampling ceased during winter of 1989–1993, gaps in the data were too wide to be filled by the moving window, so I interpolated the data linearly across the gaps. Then the data were exported as individual files for each taxon, each of which contained the year, julian day, and for each box either the mean log biomass or the proportion of zeros. The data are then used in the IBM by taking antilogs, then sampling from a binomial distribution using the proportion of zeros to get the probability of getting a zero; if a uniform (0,1) random number exceeds this probability the result is set to zero.

Small zooplankton (collected with a pump, EMP only): The pump sample is filtered through a 154 µm mesh before it is concentrated at 45 µm for counting, so the values (m^{-3}) calculated from counts of the net and the pump samples are considered additive. The pump collects mostly rotifers, nauplii, and small copepods such as *Limnoithona* spp. Rotifers have been uncommon in the estuary since the late 1980s, and have not been reported as common in delta smelt. Copepod nauplii have not been consistently identified to genus or species, so they formed a discrete group, and *Limnoithona* another one from the small zooplankton.

Data for small zooplankton were analyzed as above, except that data were available for fewer dates, so a 60-day window was used for interpolation. Also no monitoring data for small zooplankton have been gathered in the Cache Slough complex, so the means of Boxes 4 and 5 were used to populate Box 12.

Feeding rate parameters and outcomes

Values of preyK and preyV (Eqn. 1, Fig. 2) were determined iteratively to achieve three objectives: 1) Set the mean proportion of maximum feeding among all samples in the calibration period (WY 1991-2001) at about 75%; 2) Prevent fish from eating prey they are not found to eat at their life stage, probably because the prey are too large or too small; and 3) Have the model fish eat prey in approximate proportion to the available data on food consumption. This was done in R by box and sampling date; because I did not weight boxes by smelt abundance this method does not constitute a calibration of food consumption within the IBM but is meant to provide a starting point for that calibration.

A data set of prey availability was constructed by selecting each sampling event during the calibration period (water years 1991–2001) when both large and small zooplankton were collected (i.e., the IEP-EMP sampling events) and calculating the biomass of each prey type from each event in each box. These excluded Box 12 because no small-zooplankton samples have been taken there, and that region has not been sampled for most of the period of record. The data for each sample date, box, and taxon included the mean natural log of biomass (zeros excluded) and the proportion of zeros. Antilogs of the means were enhanced by half of the variance ($0.84^2/2$) to account for the skewness in the lognormal distribution. Then the antilog values were multiplied by the proportion of zeros in the raw data. The values resulting from this calculation were used to calculate feeding by each life stage for months when these life stages were abundant in the model: March–May (larvae), April–July (post-larvae), July–December (juveniles), and January–March (adults).

First the preyV values were set to limit the range of prey for each life stage. We assumed that larvae could not feed successfully on adult calanoid copepods or any cladocerans (Nobriga 2002), that post-larvae could consume all prey except the large cladoceran *Daphnia*, and juveniles could consume all prey except copepod nauplii (Nobriga 2002). We assumed that adults also would not eat nauplii, and would rarely detect the small, cryptic cyclopoid copepod *Limnoithona* spp., which in any case is not very abundant before March when abundance of adult smelt declines.

Then the preyK values were adjusted to make the proportions of each taxon in the diets reflect the proportions shown in the literature on delta smelt diets by life stage (Feyrer et al. 1993, Lott 1998, Nobriga 2002, Baxter et al. 2010, and Slater and Baxter 2014). Most of these data are aggregated across dates and locations, and some are for gut contents only and corresponding plankton counts are not reported. Therefore comparisons of model output to information in these reports were qualitative, and done by life stage within their season of abundance since both the life stages and the composition of the ambient zooplankton have a strong seasonal pattern.

Initial guesses at the preyK values were based on the calibration of Rose et al. (2013a), expanded to include the additional taxa. I adjusted these iteratively by changing values by small amounts, then comparing proportional consumption with literature reports. Then the preyK values with non-zero preyV values for each stage (i.e, the prey that could be consumed by the stage) were adjusted up or down proportionally until the mean value of the proportion of maximum consumption (C_j' in Eqn. 1) was close to 0.75. These preyK

values (Table 4) were then used in a similar analysis for the validation period, and as initial input to the IBM.

The realized consumption rates for the calibration period were close to the target but with wide ranges, especially for the younger stages (Table 5, Fig. 3). The statistics changed only slightly when the zooplankton data from the validation period was used (Table 5). Differences among boxes are apparent, likely a result of spatial patterns in abundance of zooplankton taxa with low preyK values; for example, values for larvae and postlarvae in Box 1 (Sacramento River) were generally lower than those in other boxes, reflecting dilution of zooplankton during the low-flow spring and early summer. Values for juveniles were generally higher in boxes 1–5 than in boxes 7-11, reflecting low abundance of adult and juvenile *P. forbesi* (see below).

To examine feeding by prey type with the selected values of preyK (Eqn. 1) I took 50 random samples from data for each life stage, sorted the samples to make adjacent samples most similar in relative prey composition, and plotted the proportions by biomass of each taxon in the available prey and in the calculated daily consumption by taxon from Eqn. 1.

Figures 4–7 show the proportions of biomass available and proportions eaten by taxon for each life stage of delta smelt, including only prey taxa with preyV=1 for a given life stage. The food available to larvae (Fig. 4) was dominated by copepod nauplii and juvenile copepods. The larvae consumed mainly juvenile copepods, primarily *P. forbesi* and other calanoids. Post-larvae are able to consume a larger array of prey types than larvae (Fig. 5); the more diverse prey field contained high proportions of copepod nauplii, juvenile and adult copepods, and some cladocerans, but the fish ate mostly juvenile copepods. The juvenile prey field was also diverse, partly because of seasonal decreases in abundance of warm-water species such as *P. forbesi* (Fig. 7). Consumption was similarly diverse though it emphasized adult copepods when they were available, especially *P. forbesi,* and total consumption was positively related to abundance of *P. forbesi.* The diets included some cladocerans and *Limnoithona* when there was little else to eat, which is consistent with gut-content data (Slater and Baxter 2014). The food availability in late winter-early spring when adults were present was much richer in cladocerans such as *Daphnia*, substantial proportions of cyclopoid copepods, and some calanoids. Adults (Fig. 7) consumed mainly adult calanoids and *Daphnia* spp., feeding at the larger end of the prey spectrum.

Feeding during the validation period resulted in similar overall consumption rates (Figs. 8–11) but prey availability and therefore feeding during the validation period was rather different because of shifts in the prey abundance patterns. For postlarvae adult copepods made up a lower proportion of the available prey, and nauplii a higher proportion, than during the calibration period, but proportions of consumption of the prey were not much different between the two periods because parameters for postlarvae were set so they did not eat many nauplii. The prey available to juveniles was very different during validation than calibration, with a much higher proportion of *Limnoithona* and *Acartiella* at the expense of *P. forbesi*. These differences are reflected in the proportions consumed. For adults the principal difference in the prey field was a lower proportion of cyclopoids and a higher proportion of calanoids, so that consumption was about split between calanoids and cladocerans.

The longer-term feeding picture (Figs. 12, 13) shows high variability in space and time. In the Delta the Sacramento box often had the lowest feeding rate, which is consistent with the low zooplankton abundance owing to dilution by river flow. The Delta does not show a consistent downward trend across life stages,

and for juveniles, feeding was more consistent and usually somewhat higher after the early 1990s than before. In Suisun Bay there were several years of very low feeding around 1990, and after that a consistently lower feeding rate for juveniles than before that period. Suisun Marsh (Box 9) was somewhat anomalous in its pattern.

Table 1. Changes made to algorithm for assigning zooplankton abundance to boxes.

Feature	Rose et al. 2013a	Changes in current version
Time frame	1995–2005	Water years 1991–2001 for calibration, 2001–2011 for validation
Number of taxa	6	10 large 2 small
Method for zero catch	e^{-8} for large zoops, e^{-4} small so data could be log transformed.	Zero-inflated method with two parameters per box & day: mean of ln(biomass) for non-zero values, and proportion zeros.
Standard deviation	Calculated and expanded as for mean	Constant value of 0.84 based on data
Number of boxes	11	12 with CSC split off from Box 1 (Sac)
Stations included	41 from 20mm, 29 EMP	51 from 20mm, 45 TNS, 41 MWT, 30 EMP + ~30 salinity-based stations/year starting 1994
EMP sampling	Monthly sampling	Before 1989 EMP sampled twice monthly but not in winter. 1990-1993 throughout year; from 1994 monthly at ~40% of stations, and added 2–4 salinity-based stations.
20mm sampling	Consistent	Box 12 1 station 1995, up to 12 by 2011
TNS sampling	NA	Started in 2005
MWT sampling	NA	Started in 2005

Table 2. Characteristics of data sources for zooplankton abundance. All programs continue, but years listed include all for which data were available in May 2017.

Source	Years	Months	Stations	Net mesh	Small zoops
Zooplankton monitoring	1972–2017	Jan–Dec*	37 (6–60)	154 μm	45–154 μm
20mm survey	1995–2016	Mar–Aug	41 (5–52)	160 um	—
Summer townet survey	2005–2015	Jun–Aug	32 (9–32)	160 um	—
Fall midwater trawl survey	2005–2015	Sep–Dec	32 (23–40)	160 um	—

* Mar–Nov in 1972–1976 and 1984–1994

Table 3. Zooplankton taxonomic groups and links to individual taxa. The size class is based on either sampling with a net (Large) or pump (Small). Species code is that used in the EMP zooplankton monitoring program. Taxon names are the species codes unless the group includes >1 code.

Size class	Species code	Genus	Species	Taxon name	Comments
Large	acartela	*Acartiella*	*sinensis*	acartela	Prominent in diets in late summer (Start 1994)
	daphnia	*Daphnia*	spp.	daphnia	Very abundant in freshwater
	eurytem	*Eurytemora*	*affinis*	eurytem	Historically abundant, still common in spring
	pdiapfor	*Pseudodiaptomus*	*forbesi*	pdiapfor	Most common prey (Starting 1989)
	pdiapjuv	*Pseudodiaptomus*	spp.	pdiapjuv	Second most abundant zoop group (Starting 1990)
	acartia	*Acartia*	spp.	othcalad	Other calanoid adults. These collectively are important prey but no particular species is that abundant in delta smelt habitat.
	diaptom	*Diaptomus*	spp.	othcalad	
	osphran	*Osphranticum*	*labronectum*	othcalad	
	pdiapeu	*Pseudodiaptomus*	*euryhalinus*	othcalad	
	pdiapmar	*Pseudodiaptomus*	*marinus*	othcalad	
	sinocal	*Sinocalanus*	*doerrii*	othcalad	
	tortanus	*Tortanus*	sp.	othcalad	
	othcalad			othcalad	
	acarjuv	*Acartia*		othcaljuv	Other calanoid copepodites - as for other calanoid adults.
	asinejuv	*Acartiella*		othcaljuv	
	diaptjuv	*Diaptomus*		othcaljuv	
	euryjuv	*Eurytemora*	*affinis*	othcaljuv	
	sinocaljuv	*Sinocalanus*		othcaljuv	
	tortjuv	*Tortanus*		othcaljuv	
	othcaljuv			othcaljuv	
	avernal	*Acanthocyclops*	*vernalis*	othcyc	Cyclopoid copepods (all stages) can be abundant at times and occur in delta smelt guts.
	oithdav	*Oithona*	*davisae*	othcyc	
	oithsim	*Oithona*	*similis*	othcyc	
	oithspp	*Oithona*	spp	othcyc	
	cycjuv			othcyc	
	othcycad			othcyc	
	bosmina	*Bosmina*	*longirostris*	othclad	Other cladocera can be abundant at times.
	ceriodap	*Ceriodaphnia*	spp.	othclad	
	diaphan	*Diaphanosoma*	spp.	othclad	
	othclado			othclad	
	harpact			other	Other taxa to complete the list
	annelid			other	
	barnnaup			other	
	chironomid			other	
	crabzoea			other	
	cumac			other	
	ostrac			other	
Small	limno	*Limnoithona*	spp.	limno	Extremely abundant, sometimes eaten (Start 1994)
	allcopnaup			copnaup	Copepod nauplii eaten by larvae

Table 4. Calibrated values of preyK, left blank where preyV is zero. At any prey concentration, higher levels of preyK correspond to lower feeding rates. These values were used in all subsequent analyses

Taxonomic group	Description	preyK values for all preyV not 0			
		Larvae	Post-larvae	Juveniles	Adults
acartela	*Acartiella sinensis* (copepod) adults		75	2	0.15
eurytem	*Eurytemora affinis* (copepod) adults		13	1	0.15
pdiapfor	*Pseudodiaptomus forbesi* (copepod) adults		5.2	0.9	1.5
othcalad	Other calanoid copepod adults		13	2	0.45
pdiapjuv	*Pseudodiaptomus forbesi* copepodites	0.3	0.5	3	1.5
othcaljuv	Other calanoid copepodites	0.3	0.5	5	1.5
limno	*Limnoithona* spp. copepods (all stages)	1.8	4.5	10	13.5
othcyc	Other cyclopoid copepods (all stages)	1.2	2.2	3	1.5
allcopnaup	Copepod nauplii (all spp.)	5.4	50		
daphnia	*Daphnia* spp. (cladocerans)		75	3	0.15
othclad	Other cladocerans		50	10	0.7
other	All other taxa		24	15	1.5

Table 5. Summary statistics for maximum proportional feeding rate (Eqn. 1) using parameters in Table 4. Parameters preyK for the calibration period were adjusted to get a mean proportional feeding rate of ~0.75. Parameters were then used for the validation period without adjustment. N is total number of samples, and all maxima were over 0.97.

	N	Mean ± SD	Median	Minimum
		Calibration period		
Larvae	420	0.75 ± 0.18	0.79	0.23
Post-larvae	597	0.75 ± 0.18	0.79	0.13
Juveniles	813	0.75 ± 0.17	0.80	0.20
Adults	268	0.76 ± 0.13	0.77	0.41
		Validation period		
Larvae	329	0.77 ± 0.17	0.83	0.26
Post-larvae	439	0.76 ± 0.16	0.79	0.16
Juveniles	659	0.71 ± 0.18	0.75	0.19
Adults	329	0.78 ± 0.12	0.79	0.49

Figure 1. Flow diagram summarizing the process for preparing zooplankton biomass data for input to the IBM. Each text box represents a process, and blue gridded boxes represent matrices or arrays of data.

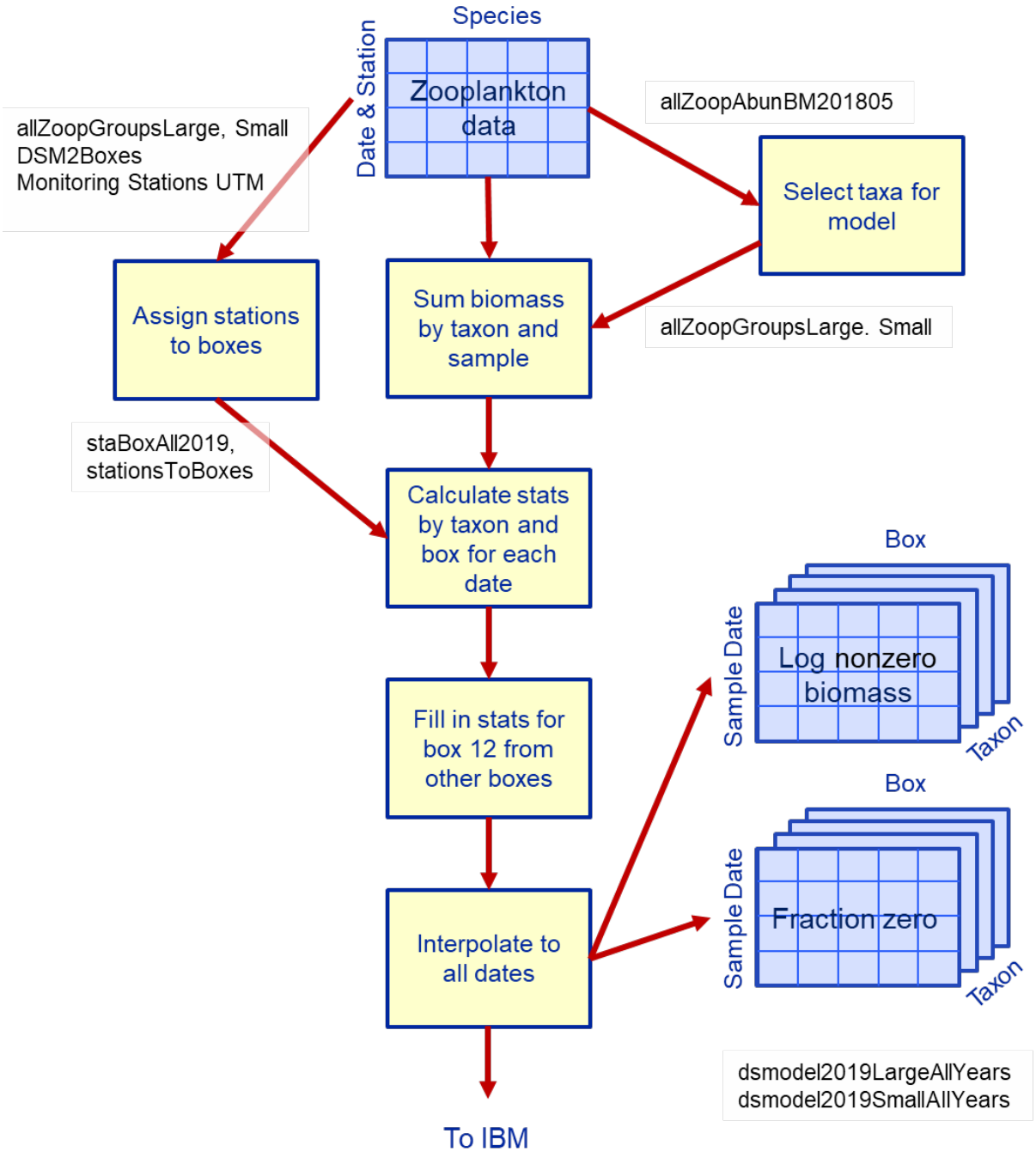


Figure 2. Flow diagram summarizing the process for determining preyV and preyK values for input to the IBM. Shapes as in Fig. 1.

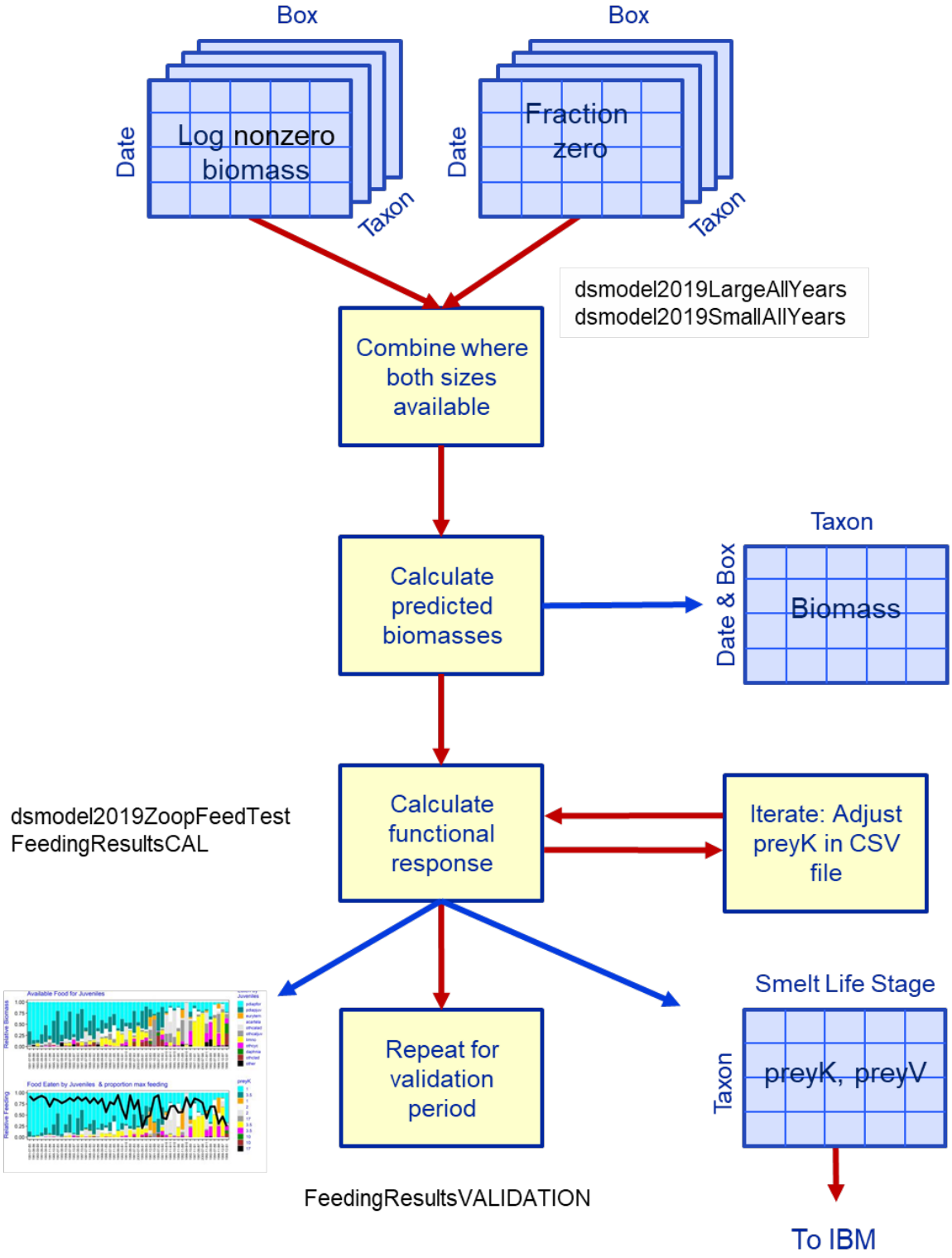


Figure 3. Boxplot of the proportion of maximum consumption (C_j' in Eqn. 1) by spatial box by delta smelt life stage for the calibration and validation periods.

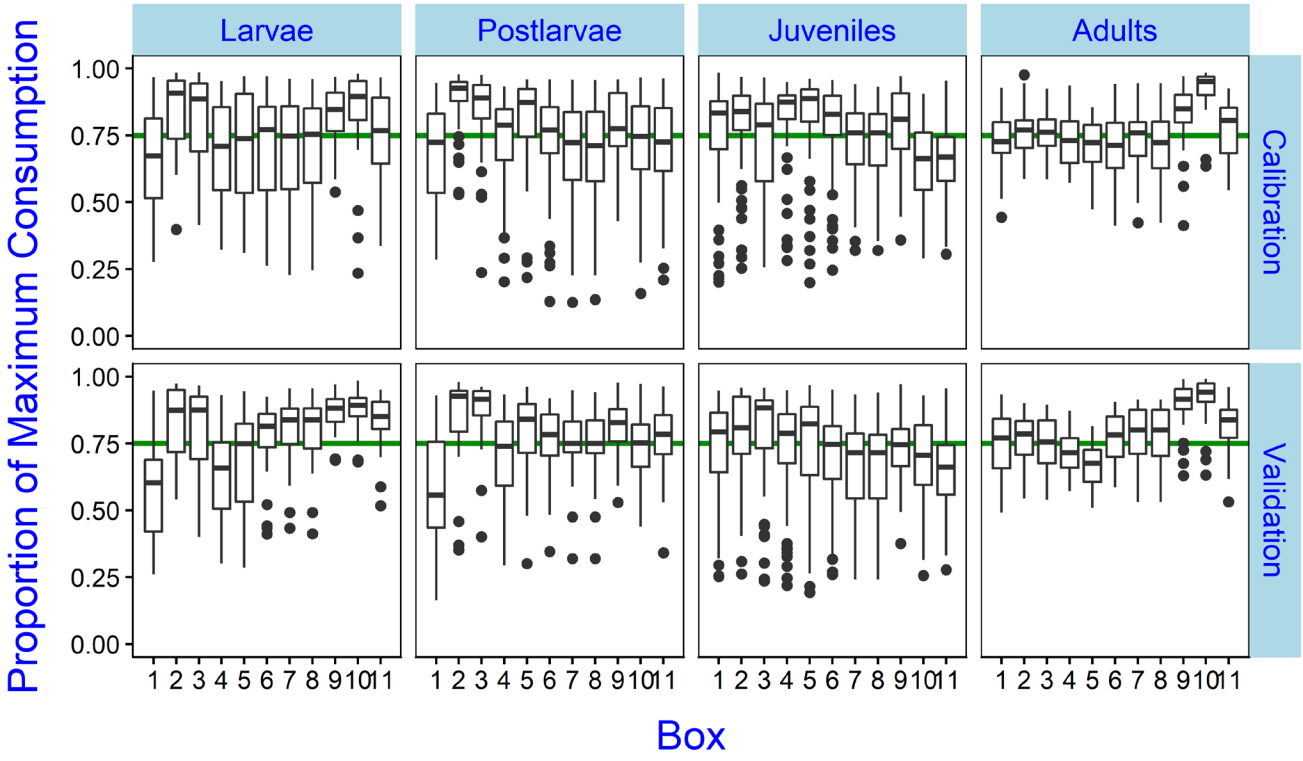


Figure 4. Calibration period. Prey available and feeding by larval delta smelt. Upper panel shows relative biomass of prey taxa and lower panel shows relative feeding on each taxon, with the heavy line indicating the proportion of maximum consumption (C_j' in Eqn. 1). Data were randomly selected from the total available (Table 5) and ordered to keep similar samples close together. Labels on the X axis denote year-month-Box. Taxon names are identified in Table 3.

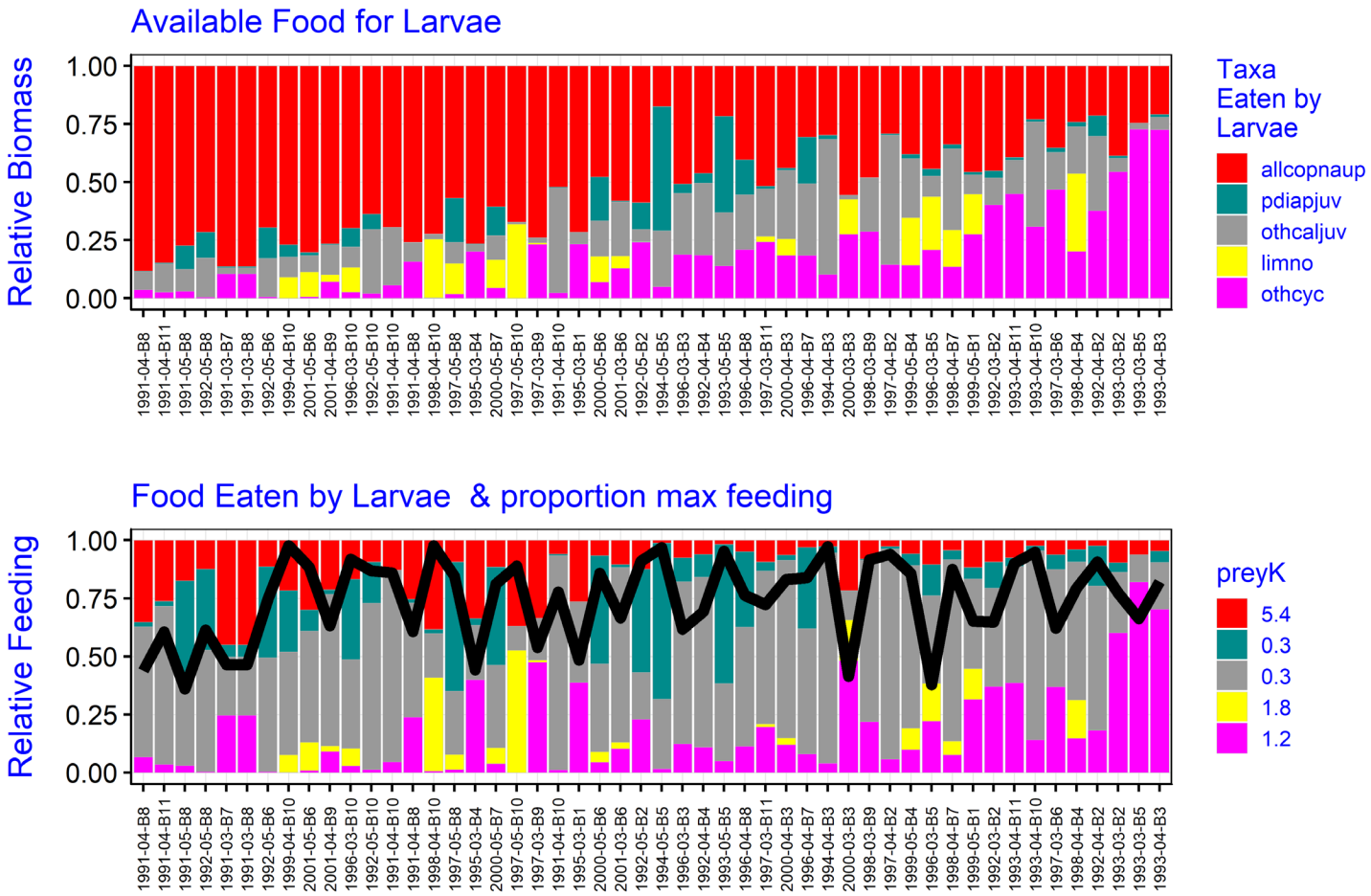


Figure 5. As in Fig. 4 for postlarvae.

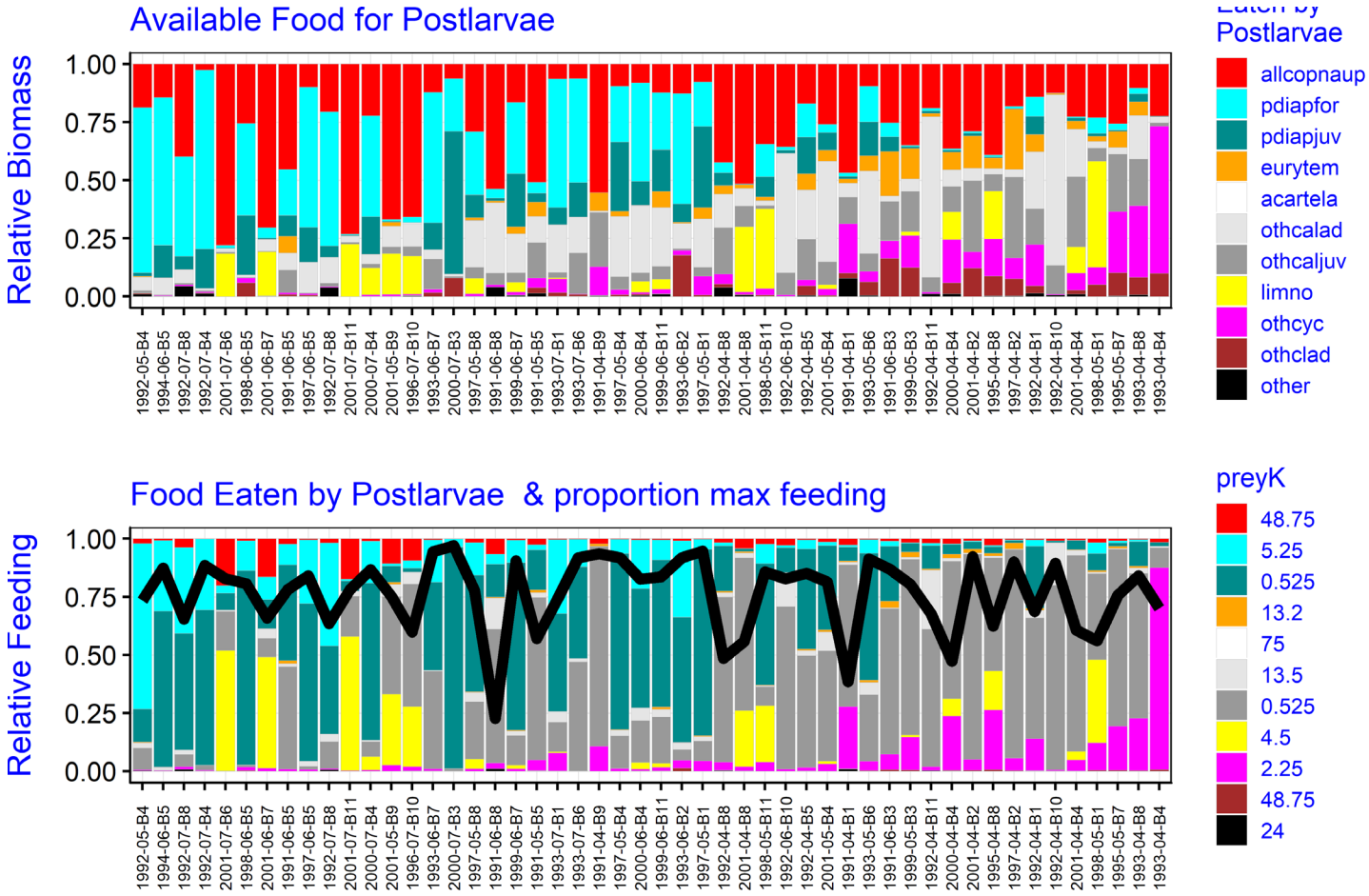


Figure 6. As in Fig. 4 for juveniles.

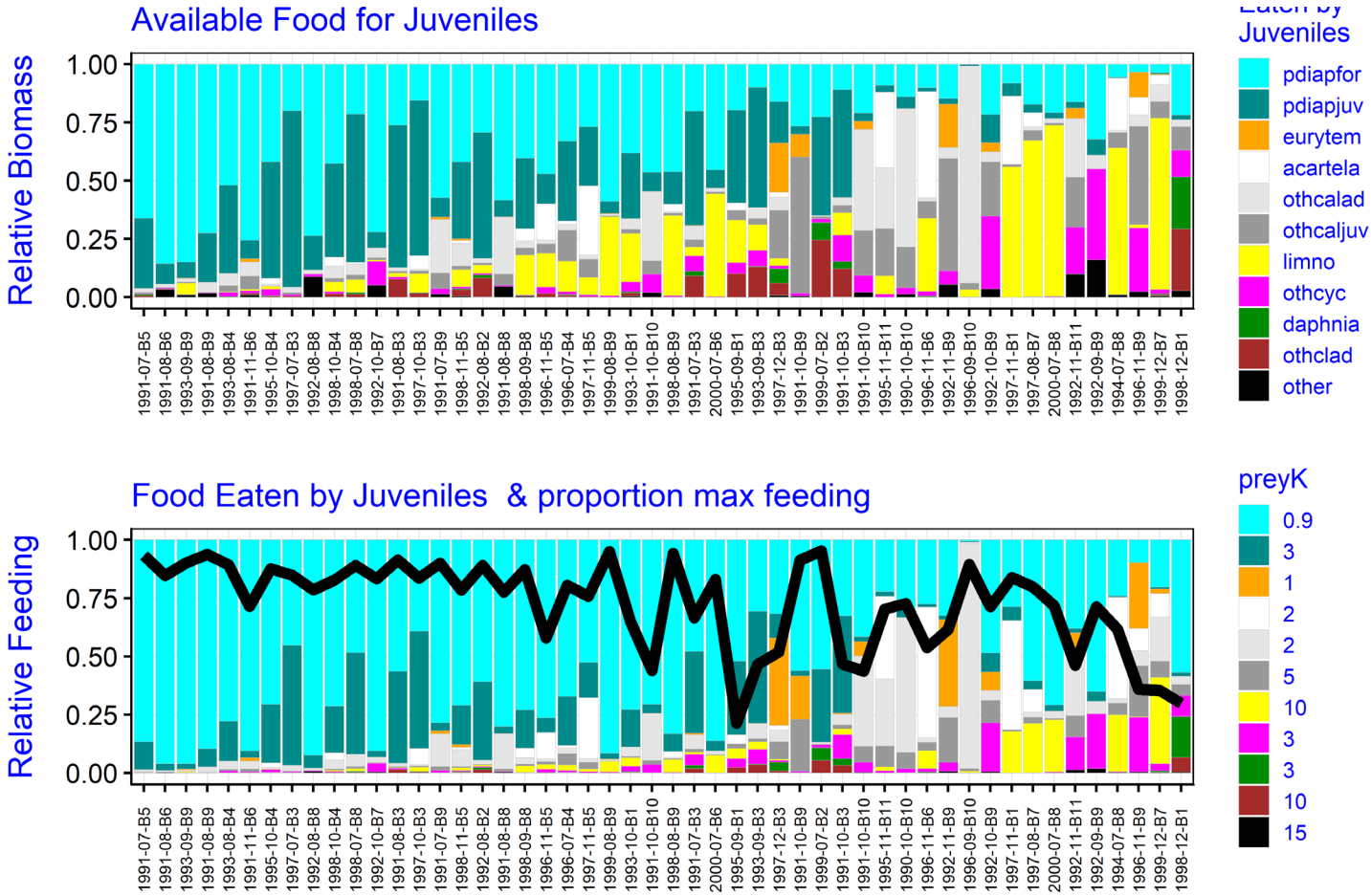


Figure 7. As in Fig. 4 for adults.

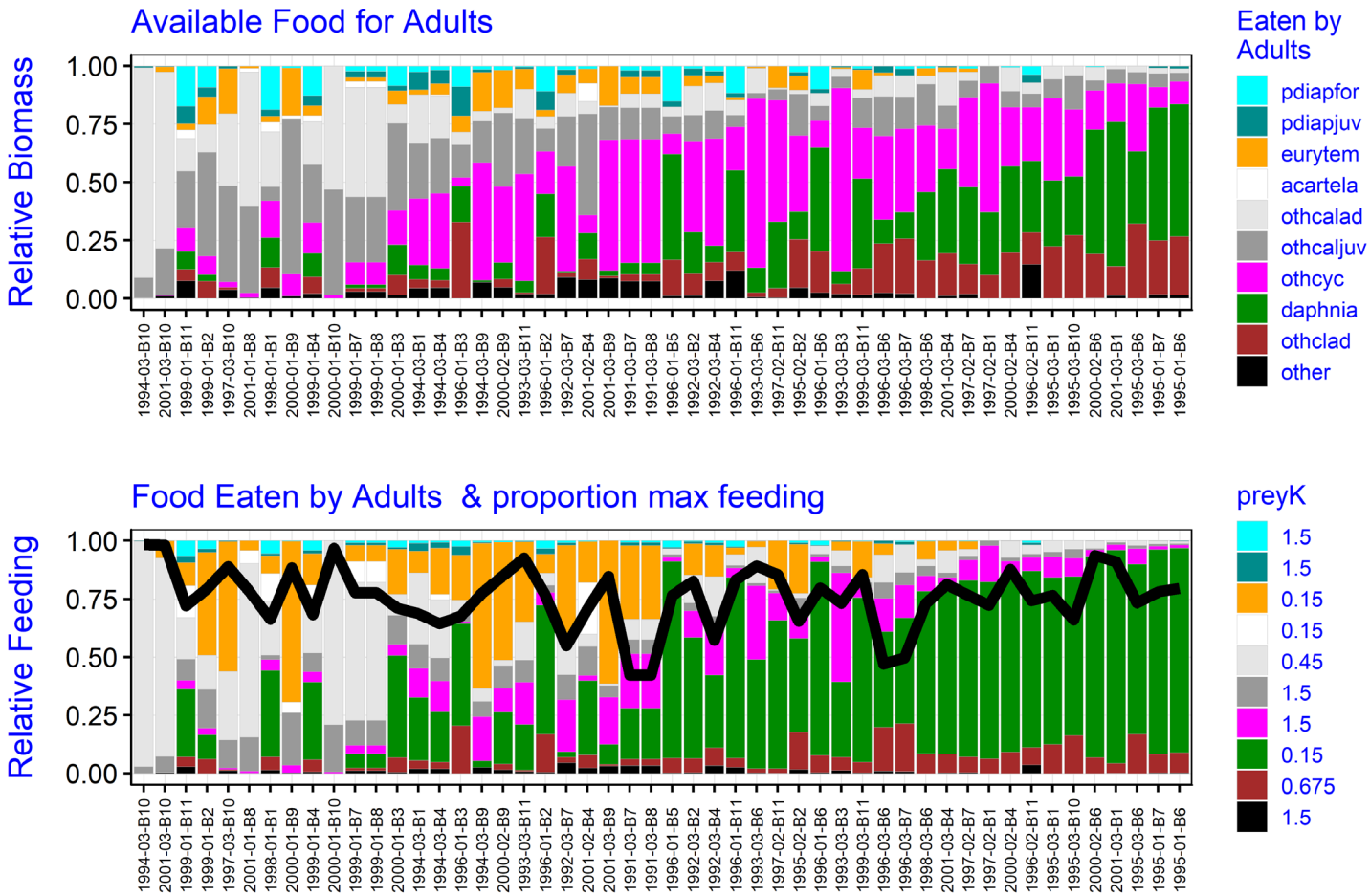


Figure 8. As in Fig. 4 for larvae in the validation period.

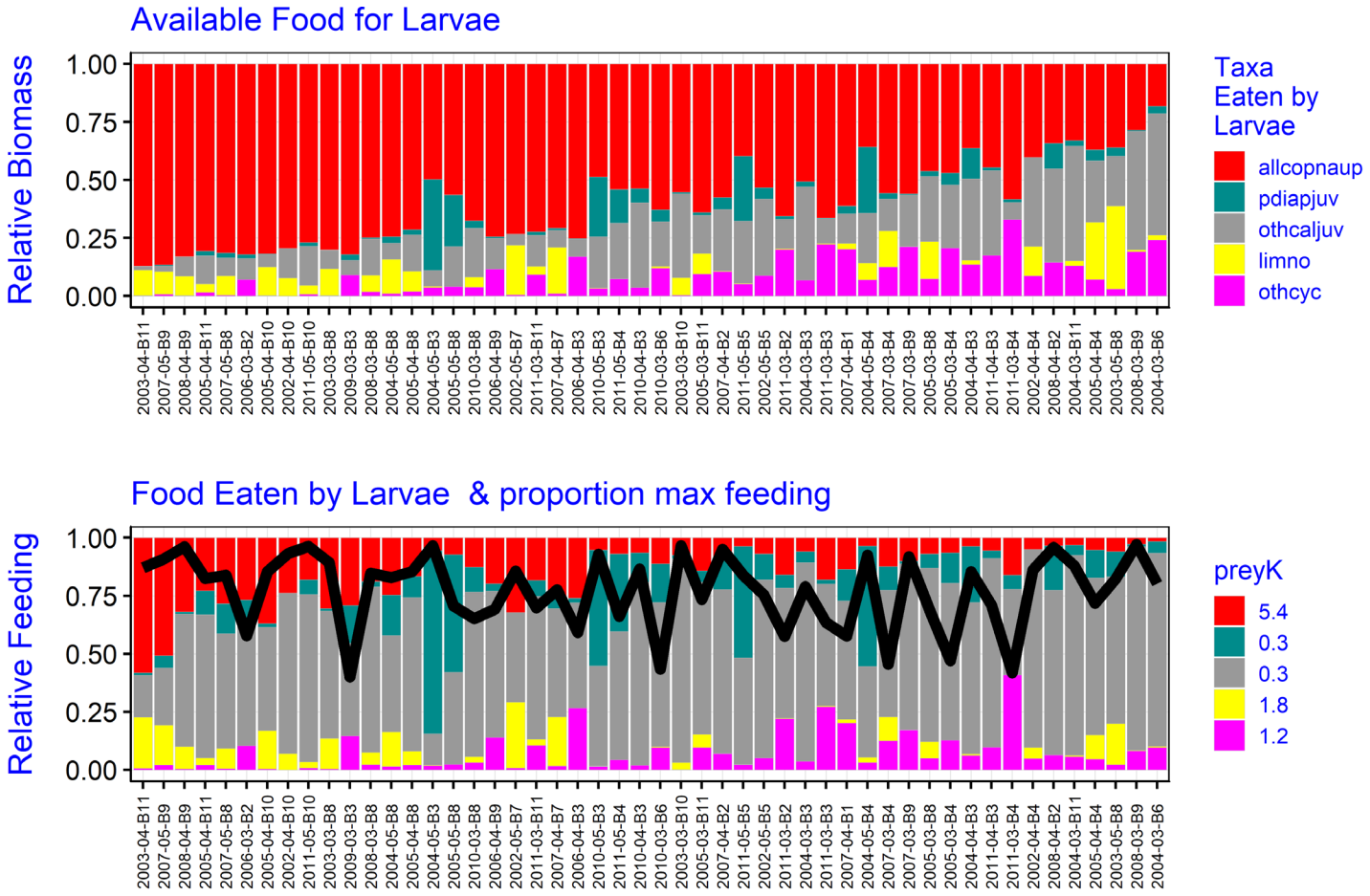


Figure 9. As in Fig. 4 for postlarvae in the validation period.

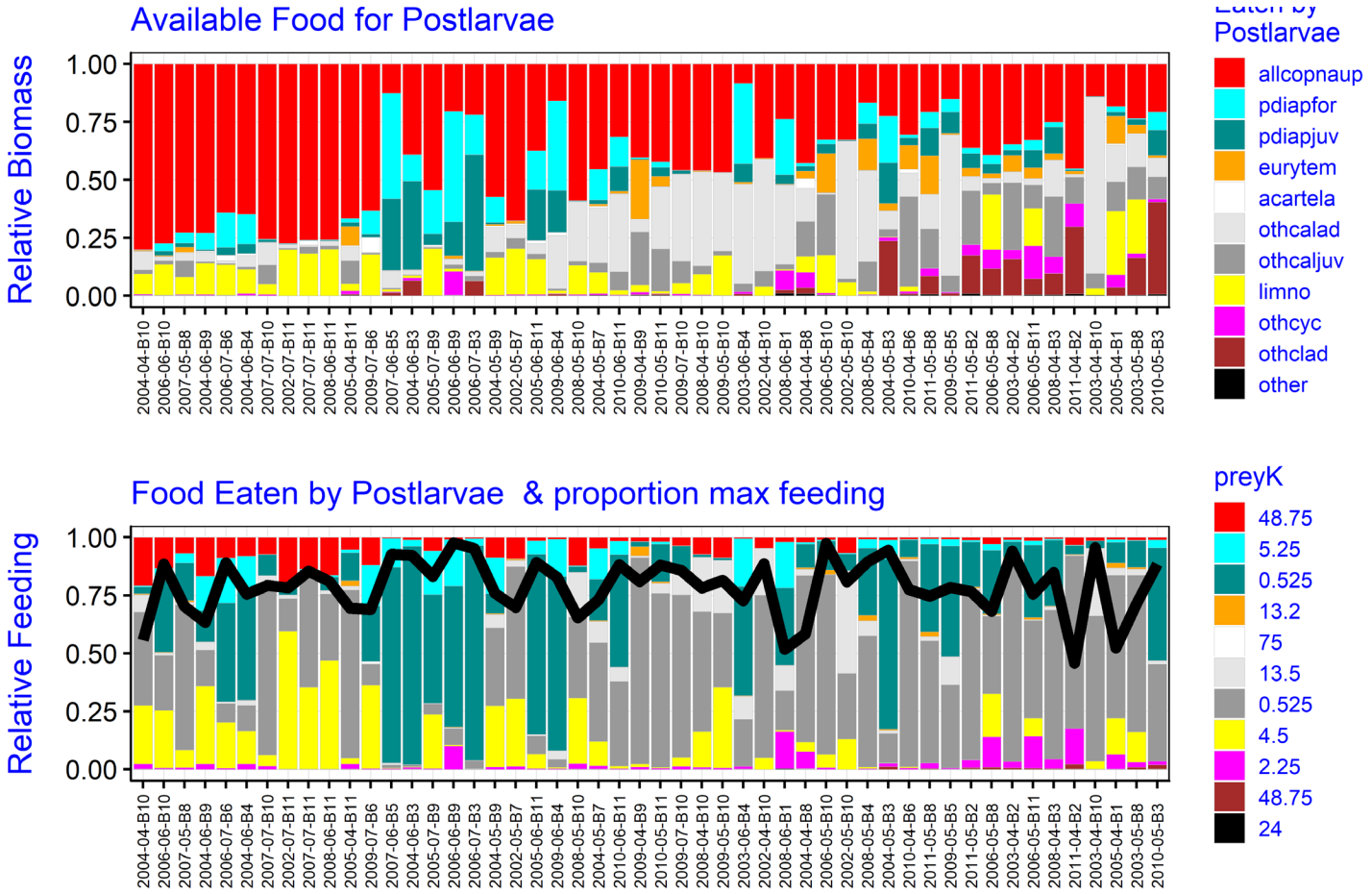


Figure 10. As in Fig. 4 for juveniles in the validation period.

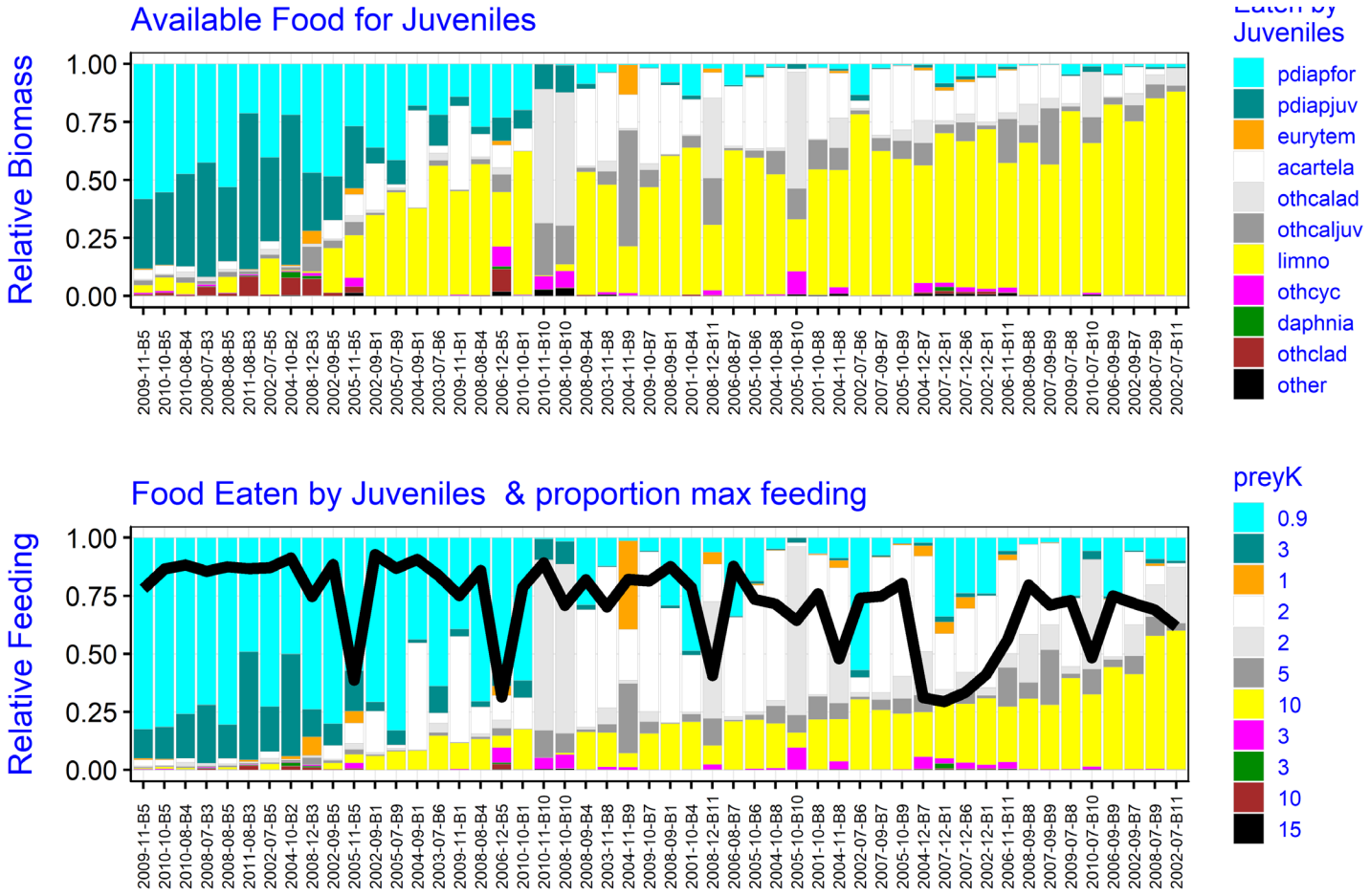


Figure 11. As in Fig. 4 for adults in the validation period.

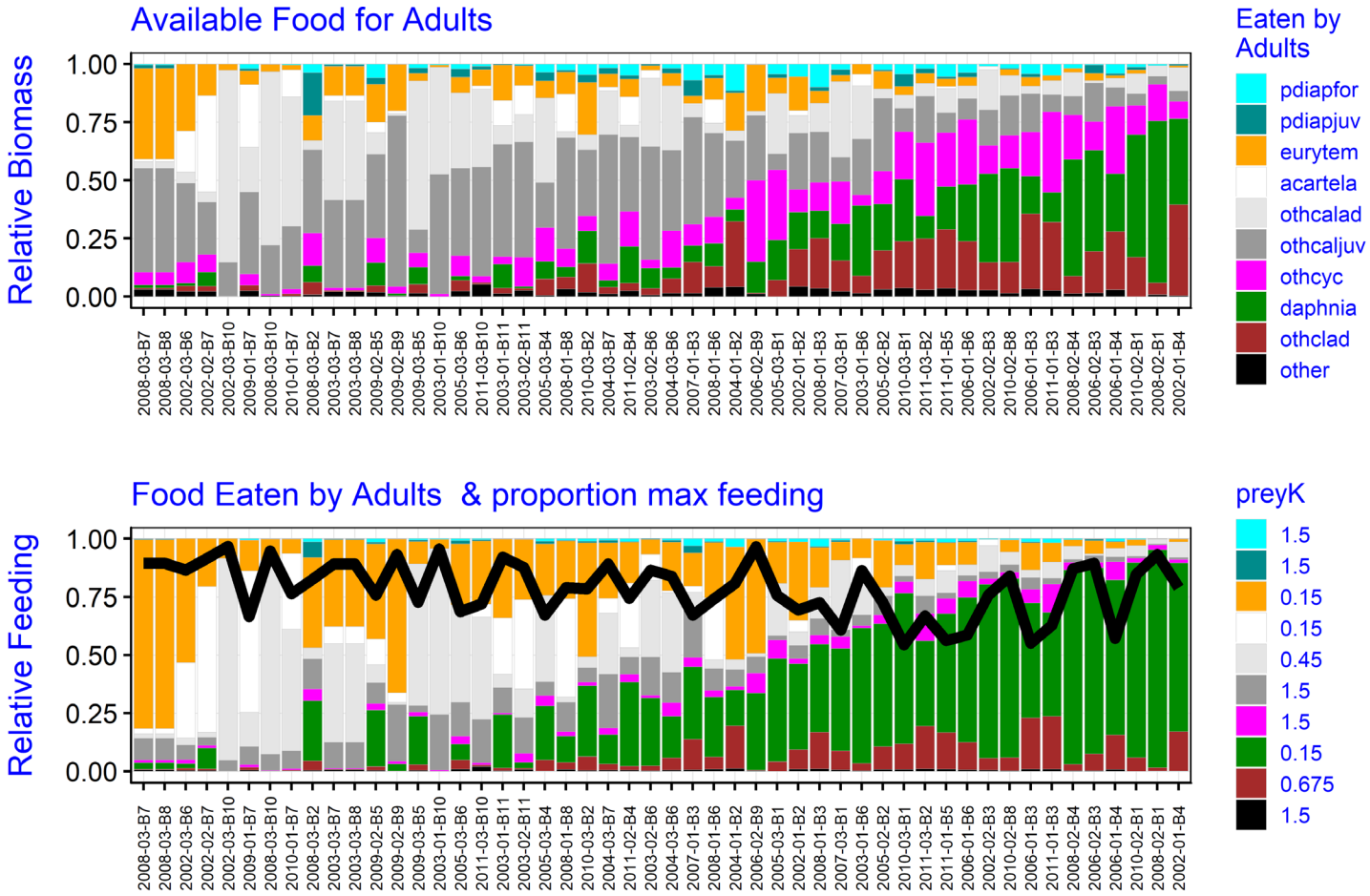


Figure 12. Fraction of maximum feeding rate vs. year by life stage for Delta (boxes 1-6) and Suisun Bay and Marsh. Values are means by year for the season of abundance of the life stage.

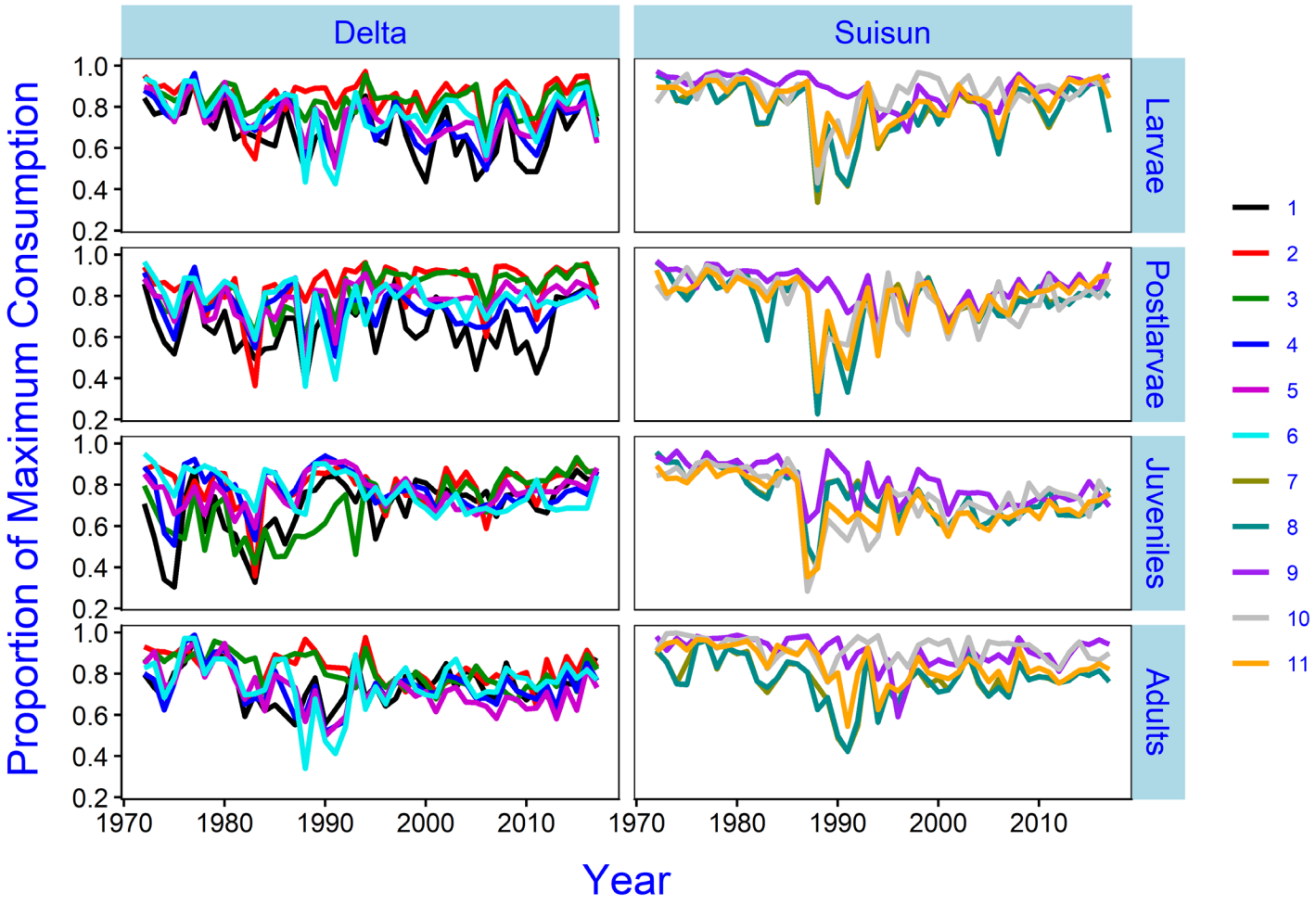
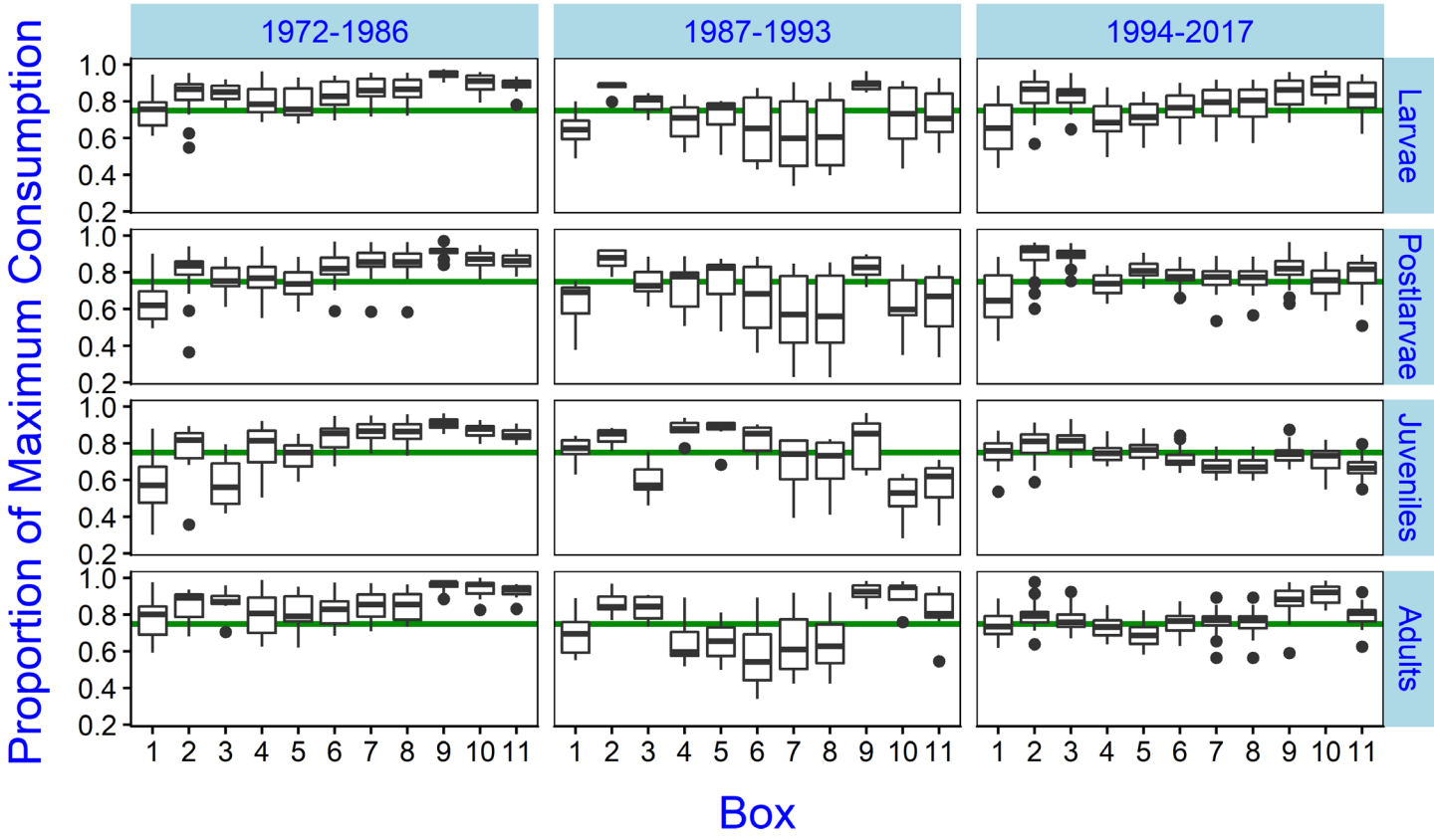


Figure 13. Data in Fig. 12 arranged as boxplots for three time periods reflecting major periods of change in the estuarine food web.



Appendix B

Missing temperature and Secchi depth data

Introduction

General linear models were developed to predict missing temperature and turbidity data. The objective of these models was prediction, not inference; therefore, covariate effects were not selected based any particular mechanistic link (e.g., spatial proximity), and the relative effects within each model were not compared. Only explanatory power and model performance was considered.

Temperature

DSM2 monthly mean $Temp_{\mathrm{yms}}$ for years *y*, 2011-2014, months *m*, and spatial strata *s* were predicted as a function of season, spatial strata, and monthly mean temperatures measured by fish monitoring programs $\widehat{Temp}_{\mathrm{yms}}$, using a general linear model. $Temp_{\mathrm{yms}}$ from DSM2 represented a water column mean, while $\widehat{Temp}_{\mathrm{yms}}$ from monitoring programs represented surface measurements. The relationship between $Temp_{\mathrm{yms}}$ and $\widehat{Temp}_{\mathrm{yms}}$ was expected to vary seasonally, with the onset of thermal stratification as water warmed, and the effect of thermal stratification was expected to vary spatially, as water depth, tidal influence, and stratification vary spatially. Factorial seasonal and strata effects accounted for this spatiotemporal variation in the model to predict missing $Temp_{\mathrm{yms}}$.

$$\text{B1. } Temp_{\mathrm{yms}} = \begin{array}{c} \beta_{Temp_0} + \beta_{Temp_1} * \widehat{Temp}_{\mathrm{yms}} + \\ \sum_{i=1}^{n.season} \beta_{Temp_{2+\mathrm{i}}} * Season_{\mathrm{m}} + \\ \sum_{j=1}^{n.st.group_{Temp}} \beta_{Temp_{2+n.season+\mathrm{j}}} * St.group_{Temp_{\mathrm{s}}} \end{array}$$

β_{Temp} represented coefficients of the general linear model, the quantity $\beta_{Temp_0} + \sum_{i=1}^{n.season} \beta_{Temp_{2+\mathrm{i}}} * Season_{\mathrm{m}} + \sum_{j=1}^{n.st.group_{Temp}} \beta_{Temp_{2+n.season+\mathrm{j}}} * St.group_{Temp_{\mathrm{s}}}$ represented a unique intercept for each *s*, and $\beta_{Temp_1} * \widehat{Temp}_{\mathrm{yms}}$ scaled fish monitoring temperatures to DSM2 temperatures. Beginning with a full model, having four seasonal effects $Season_{\mathrm{m}}$, backwards selection was used to combine seasons until *n.season* groupings remained, with coefficient p-values < 0.05. After selecting seasonal effects, the same process was used to eliminate strata-specific effects. Beginning with a full model, having 12 strata-specific effects $St.group_{\mathrm{s}}$, backwards selection was used to combine strata until $n.st.group_{Temp}$ groupings remained, with coefficient p-values < 0.05.

Secchi depth

Missing Secchi depth measurements $Secchi_{\text{yms}}$ were also predicted from a general linear model. Since turbidity stratification was not expected to occur and no secondary measurements of $Secchi_{\text{yms}}$ were available from independent sources, missing $Secchi_{\text{yms}}$ were predicted using the remaining measurements in other strata.

B2. $Secchi_{\text{ymi}} = \beta_{Secchi_0} + \sum_{j=1}^{n.st_{Secchi_\text{i}}} \beta_{Secchi_\text{j}} * Secchi_{\text{ym}(St.group_{Secchi_{\text{ij}}})}$,

for *i* in the set [Yolo, East Delta, Northeast Suisun]. Backwards model selection was used to eliminate strata-specific effects β_{Secchi} until $n.st_{Secchi}$ strata effects remained, with coefficient p-values < 0.05.

Models B1 and B2 were fit using the glm() function in R (R 2018), and the best models indicated by model selection were evaluated graphically using general diagnostic plots, residual, q-q norm, standardized residual, and influence, for violations of basic linear model assumptions. Residuals were expected to be randomly distributed around zero. qq-norm plots were expected to be relatively linear, and the slope of standardized residuals was expected to be zero.

Results

The best model of $Temp_{\text{yms}}$ included separate spring and summer effects, but winter and fall were grouped (Table B1). Most strata were grouped, but South Delta, Confluence, and Southwest Suisun strata each received unique coefficients, resulting in unique intercepts for these three strata. Diagnostics did not indicate severe violations of general linear model assumptions. Model residuals appeared to be normally distributed, and no extreme outliers or leverage points were detected (Fig. B1). The model explained 95.8% of the variation in the $Temp_{\text{yms}}$ data, indicating high explanatory power.

All missing $Secchi_{\text{yms}}$ data were from years prior to 2011. Of 240 year-month combinations, 30 $Secchi_{\text{ym(s=Yolo)}}$ were missing, 7 $Secchi_{\text{ym(s=East Delta)}}$ were missing, and 1 $Secchi_{\text{ym(s=Northeast Suisun)}}$ were missing. The month of August was especially problematic for $Secchi_{\text{ym(s=Yolo)}}$, with zero samples from 1998 to 2010.

The best model to fill in missing $Secchi_{\text{yms}}$ varied by stratum (Table B2). The best model to predict $Secchi_{\text{ym(s=Yolo)}}$ included Sacramento, South Delta, Lower Sacramento, and Lower San Joaquin effects. Some $Secchi_{\text{ym(s=East Delta)}}$ were missing in the same time periods as the missing Yolo data, so East Delta data were not used to predict Yolo Secchi depth. The best model to predict $Secchi_{\text{ym(s=East Delta)}}$ included Sacramento, South Delta, Southeast Suisun, and Southwest Suisun effects, and the best model to predict $Secchi_{\text{ym(s=Northeast Suisun)}}$ included Sacramento, Southeast Suisun, Northeast Suisun, and Northwest Suisun effects. Diagnostics revealed exactly one extreme outlier in the datasets to predict $Secchi_{\text{yms}}$ for each stratum, so these outliers were removed prior to fitting the final model. After removing outliers, most residuals appeared to be normally distributed (Figs. B2-B4), and no high leverage points were detected. Residuals were somewhat

higher at the highest predicted values of greater than 175 cm Secchi depth, suggesting that a log-linear model might fit the data better. Modeled effects on delta smelt did not vary at Secchi depths greater than 84 cm, so an improved model of the highest Secchi depths was not expected to change the IBMR dynamics.

The explanatory power of models to predict $Secchi_{\mathrm{yms}}$ varied, with East Delta and Northeast Suisun models explaining 79.5% and 82.7% of variation, respectively. The best model to predict missing $Secchi_{\mathrm{ym(s=Yolo)}}$ explained only 61.6% of variation. Many of the Yolo observations, or y values, were limited to a single replicate sample per month, versus and average of 30 Secchi depth samples within other year-month-strata combinations. This limitation may have been a source of observation error in the Yolo model.

Discussion

The models developed here to predict missing $Temp_{\mathrm{yms}}$ and $Secchi_{\mathrm{yms}}$ leave much room for improvement.

-The best way to improve the estimation of $Temp_{\mathrm{yms}}$ data used by IBMR would be to complete DSM2 simulations for the entire 1995-2014 time series.

-Backwards selection using coefficient p-values was used to identify the best $Temp_{\mathrm{yms}}$ and $Secchi_{\mathrm{yms}}$ general linear models, but a comparison of all possible models, using AIC would be a superior method. The primary advantage would be increasing the number of strata groupings to consider. The set of all possible strata combinations is very large. For example, just consideration of all 1- and 2-strata groupings would result in a total of 78 candidate models to consider.

-Replacing the linear models with a more flexible functional form, such as a generalized additive model (GAM), could improve predictions by accounting for non-linear relationships.

-Some hydrodynamic models of the San Francisco Estuary are capable of simulating turbidity fields. Application of these models to estimate the IBMR $Secchi_{\mathrm{yms}}$ dataset would be especially useful for the Yolo spatial strata, which was limited by a low number of within-month (replicate) samples and a larger number of unsampled year-month combinations, relative to other strata.

References

R Core Team. 2018. R version 3.5.0: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Table B1. Results from the best model to predict DSM2 mean water column temperatures from fish monitoring measurements of surface temperature $\widehat{Temp}$.

Coefficient	Estimate	p-value
Intercept	1.370	$< 2x10^{-16}$
Spring	0.529	$< 2x10^{-16}$
Summer	1.649	$< 2x10^{-16}$
Southwest Suisun	-0.197	0.002
South Delta	-0.579	$< 2x10^{-16}$
Confluence	-0.280	$1.1x10^{-5}$
$\widehat{Temp}$	0.891	$< 2x10^{-16}$
% null deviance explained	95.8%	

Table B2. Results from the best models to predict missing Secchi depth data in Yolo, East Delta, and Northeast Suisun spatial strata from the remaining data in other strata.

	Yolo		East Delta		Northeast Suisun	
Coefficient	Estimate	p-value	Estimate	p-value	Estimate	p-value
Intercept	-10.219	0.0013	-7.967	0.0075	-2.303	0.0054
Sacramento	0.219	2.8×10^{-9}	0.804	$< 2 \times 10^{-16}$	-0.041	1.57×10^{-5}
South Delta	0.310	7.6×10^{-10}	0.411	1.4×10^{-14}	--	--
East Delta	--	--	--	--	--	--
Lower Sacramento	0.545	2.1×10^{-9}	--	--	--	--
Lower San Joaquin	-0.183	0.011	--	--	--	--
Confluence	--	--	--	--	--	--
Southeast Suisun	--	--	0.227	0.00011	0.269	$< 2 \times 10^{-16}$
Northeast Suisun	--	--	--	--	0.309	3.4×10^{-15}
Suisun Marsh	--	--	--	--	--	--
Southwest Suisun	--	--	0.227	0.041	--	--
Northwest Suisun	--	--	--	--	0.462	$< 2 \times 10^{-16}$
% null deviance explained	61.6%		79.5%		82.7%	

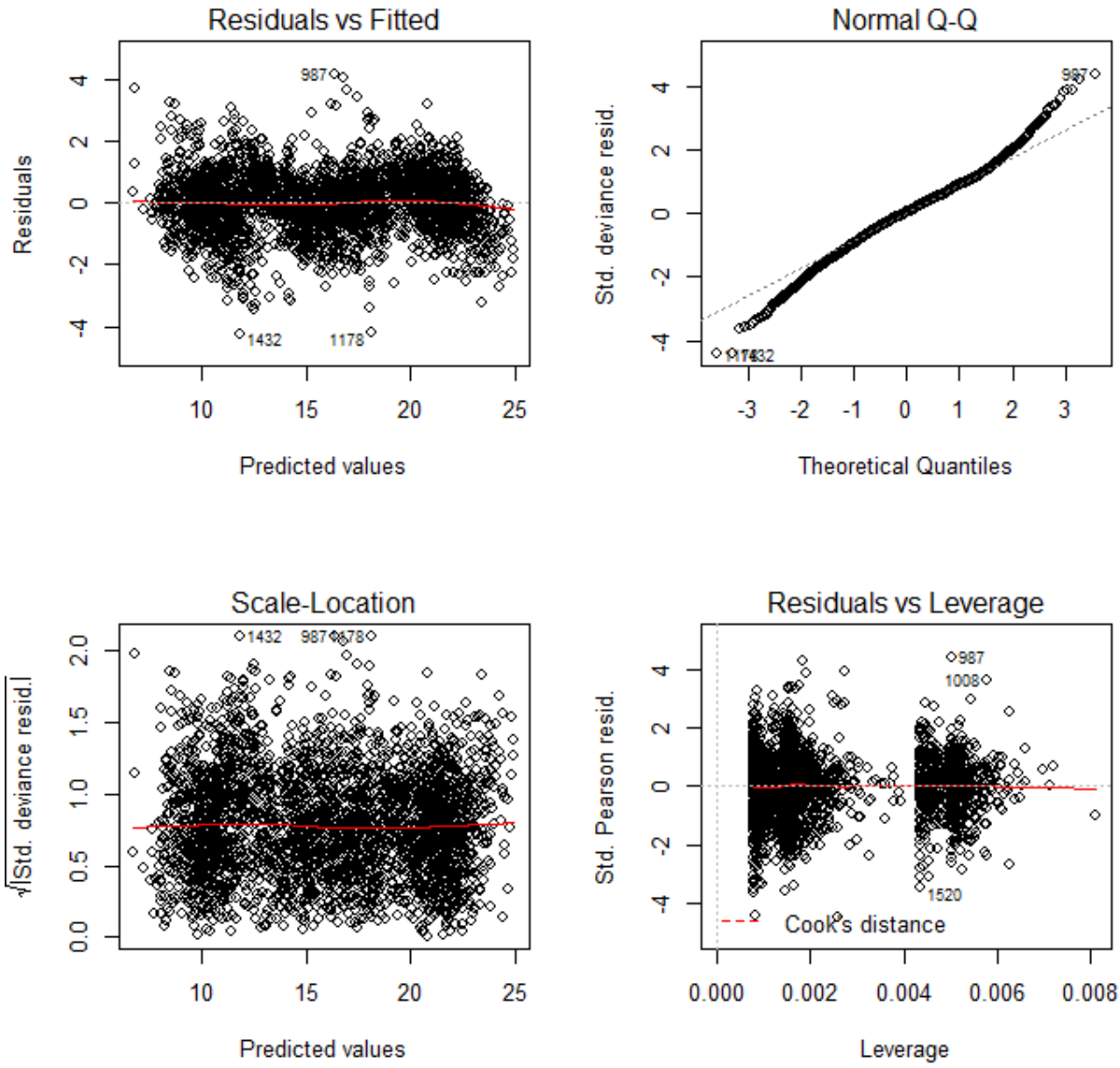


Figure B1. Diagnostic plots for the best model to predict DSM2 mean water column temperatures from fish monitoring measurements of surface temperatures.

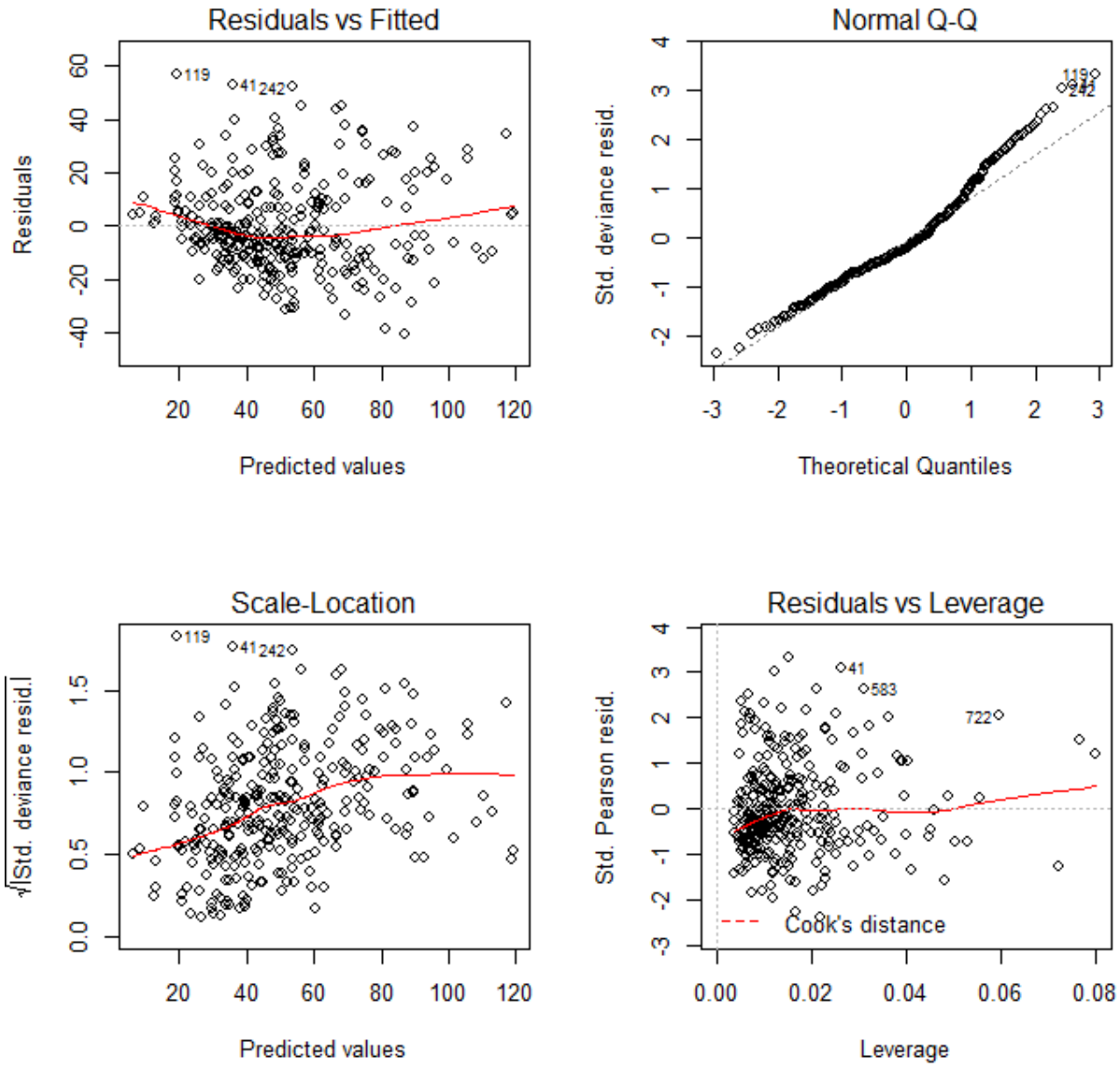


Figure B2. Diagnostic plots for the best model to predict missing Yolo strata Secchi data from data in the remaining spatial strata during the same time period.

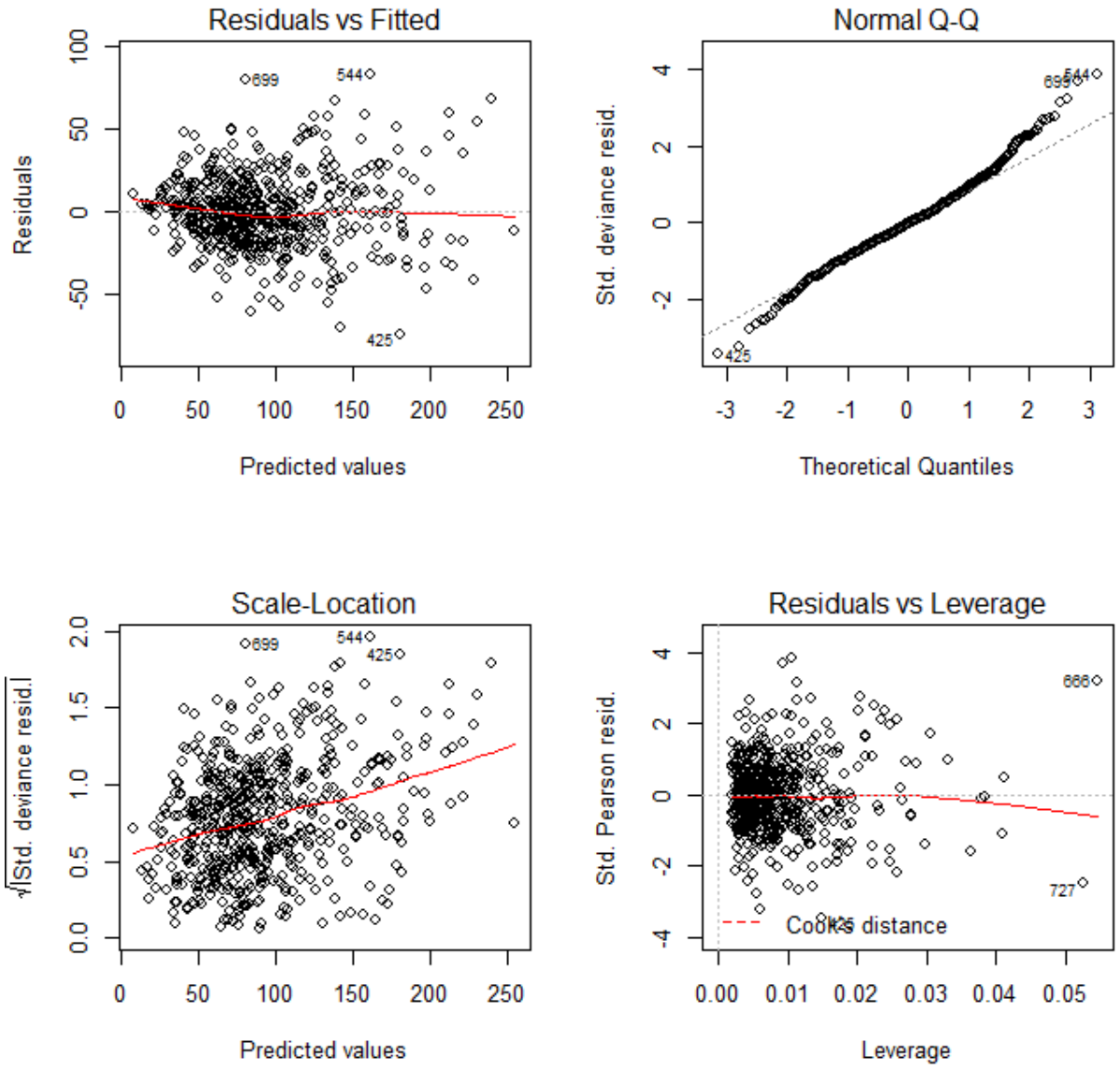


Figure B3. Diagnostic plots for the best model to predict missing East Delta strata Secchi data from data in the remaining spatial strata during the same time period.

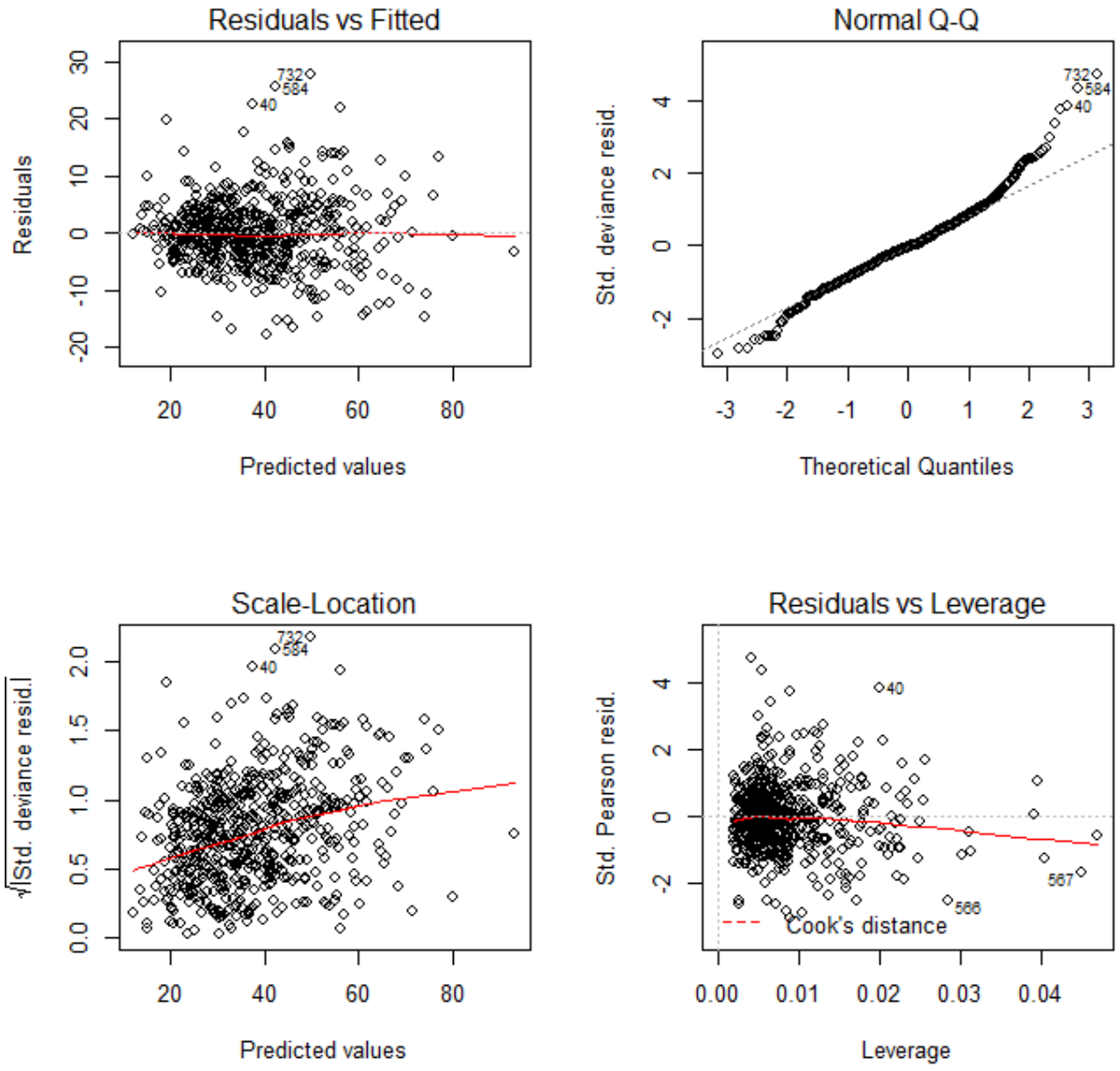


Figure B4. Diagnostic plots for the best model to predict missing Northeast Suisun strata Secchi data from data in the remaining spatial strata during the same time period.

Appendix C

Supplementary figures

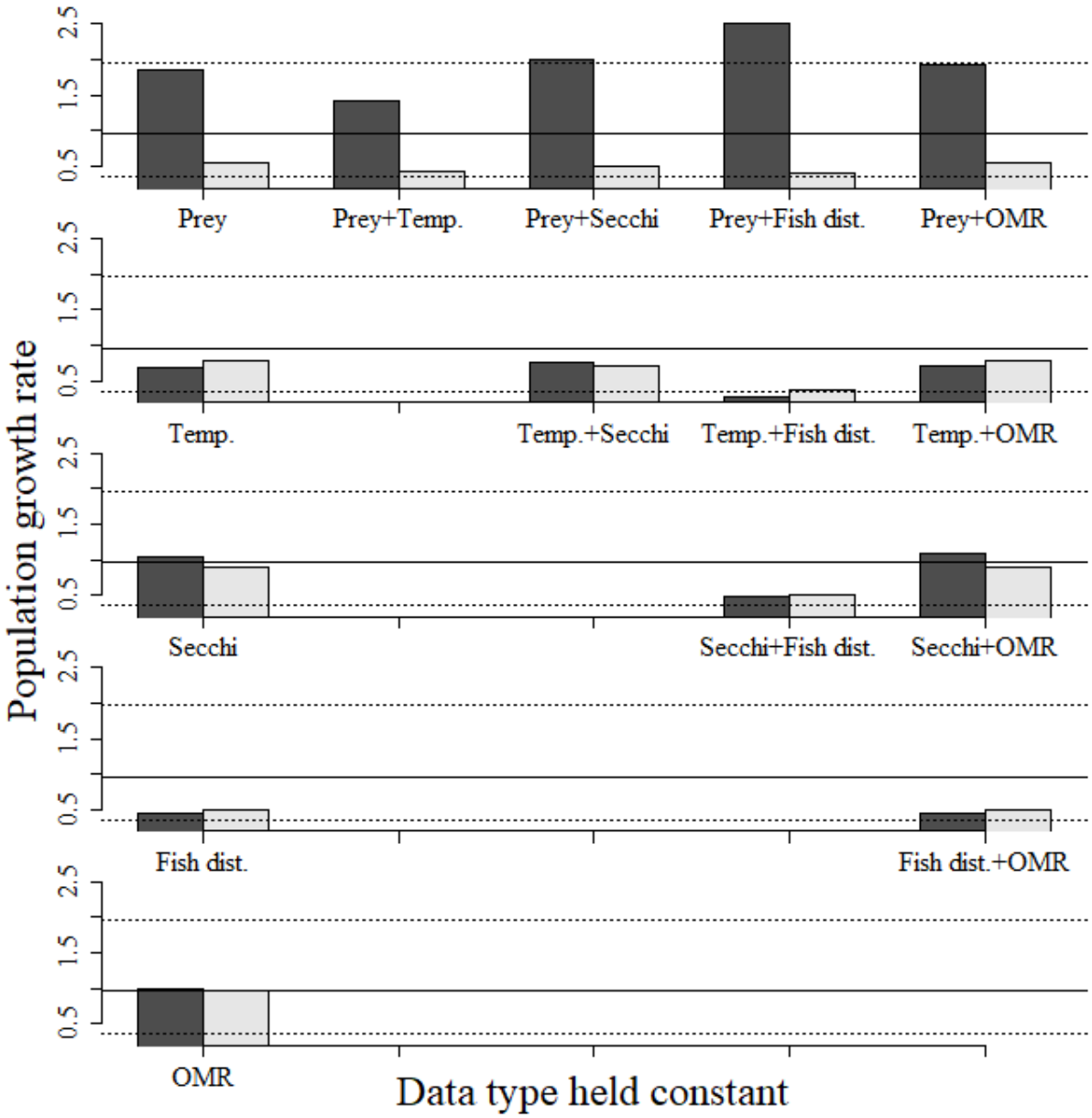


Figure C1. Results of sensitivity analysis, when South Delta distributions were modeled with fish survey observations. Bars show the geometric mean of population growth rates, when one or two data types were fixed at the values measured in high growth year 1998 (dark bars) or low growth year 2004 (light bars). Data types were prey density, temperature (Temp.), Secchi depth, observed delta smelt spatial distribution (Fish dist.), and Old and Middle River flow (OMR). Horizontal reference lines indicate the mean λ_{AB} when all data was allowed to vary at the values measured in 1995-2014 (solid line) and the maximum and minimum λ_{AB} simulated in 1998 and 2004 when all data was allowed to vary (dotted lines).

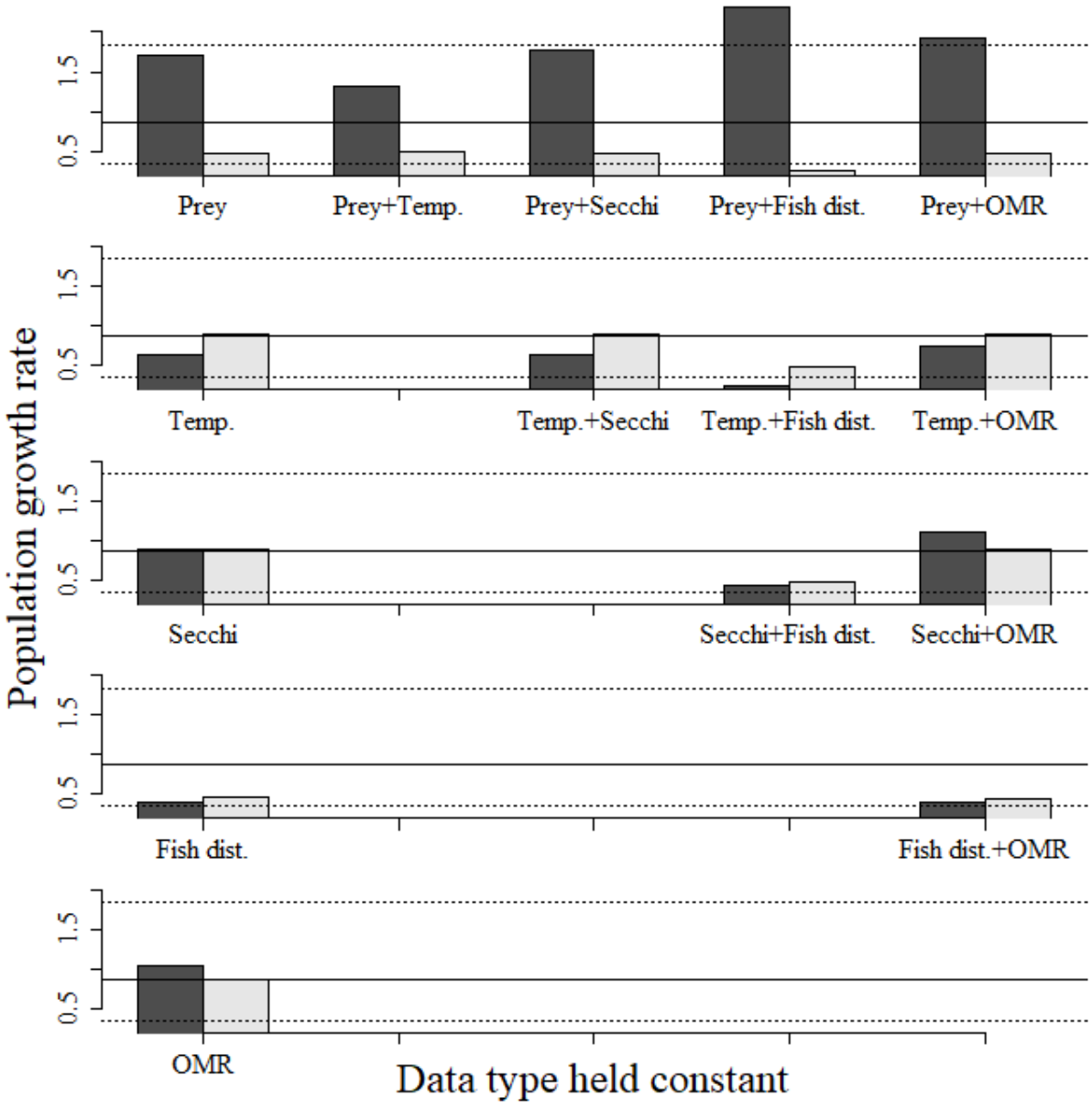


Figure C2. Results of sensitivity analysis, when no turbidity-mortality relationship was modeled. Bars show the geometric mean of population growth rates, when one or two data types were fixed at the values measured in high growth year 1998 (dark bars) or low growth year 2004 (light bars). Data types were prey density, temperature (Temp.), Secchi depth, observed delta smelt spatial distribution (Fish dist.), and Old and Middle River flow (OMR). Horizontal reference lines indicate the mean λ_{AB} when all data was allowed to vary at the values measured in 1995-2014 (solid line) and the maximum and minimum λ_{AB} simulated in 1998 and 2004 when all data was allowed to vary (dotted lines).